INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB RD., ANN ARBOR, MI 48106

8212847

Klein, Sharon Michelle

SYNTACTIC THEORY AND THE DEVELOPING GRAMMAR:
REESTABLISHING THE RELATIONSHIP BETWEEN LINGUISTIC THEORY
AND DATA FROM LANGUAGE ACQUISITION

University of California, Los Angeles PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

UNIVERSITY OF CALIFORNIA

Los Angeles

Syntactic Theory and the Developing Grammar: Reestablishing the Relationship between Linguistic Theory and Data from Language Acquisition

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Linguistics

by

Sharon Michelle Klein

1982

The dissertation of Sharon Michelle Klein is approved.

Peter W. Culicover

Joseph E. Emonds

Kenneth Wexler

Stephen R. Anderson, Committee Chair

University of California, Los Angeles

1982

To Rachael, of course, whose arrival made all this work worthwhile, and put it in perspective, and to her Dad, whose strength and push made finishing it a reality.

TABLE OF CONTENTS

Chapter

Chapter

ACKNOWLEDGMENTS

It is often the case that in works requiring the depth and breadth that a dissertation demands, the individual whose name appears below the title must ultimately bear the responsibility for all of the contents, but has benefited greatly from the ideas, insights, criticisms, suggestions and support of many others. This volume is a paradigm case of that generalization, and in these few paragraphs I wish to mention and thank those others.

No one could ask for a better committee. Peter Culicover was always willing to talk about the work, and was often able to present the counterargument that seemed devastating, but generally caused me to think things through in order to come up (in most cases) with a satisfactory response. I'm grateful too to Ken Wexler who joined the committee late, but whose contribution nonetheless has been substantive, and whose clear thinking I respect. Joe Emonds I thank for all of his ideas, his theoretical élan, his enthusiasm and confidence, his teaching and his thinking; all of which he shares freely. As a chair, finally, Stephen Anderson has been truly exceptional. Exacting and consistently insightful, he became a "presence" as I worked--often I could hear what I supposed his comments would be, and sought to respond to them. Many times a probing, argumentative session culminated in a new perspective, and enhanced, hence my views of many issues. No less important than his role as advisor has been his official role as

committee chair. On more than one occasion he deftly cut through administrative tangles that had seemed unresolvable. One whose participation has equalled that of an active member, but whose name does not appear because geography kept him from attending the defense is Henry Hamburger. Our conversations about language acquisition always provided ideas and stimulated my thinking. I value his contributions highly.

Friends and fellow students played a significant role in helping me manage to finish this work as well as in enhancing it by sharing possible hypotheses, offering criticism, and sometimes even agreeing with me. Among these I want to thank especially Julia Horvath, Nina Hyams, Wendy Wilkins, Randy Hendrick and Gerry Delahunty. I've been very lucky to know all of these people, benefiting from their exceptional intelligence, clear thinking, and good ideas. Most of all, though, I feel fortunate to be able to call them friends.

My family, needless to say, deserves substantial credit. Rachael's Grandma and Grandpa Klein came through too many times to count, as babysitters, offering all kinds of support. My parents too have always "been there" with support of all kinds, including often-needed "nudges" in the right direction. Keith gave me confidence that I could in fact do this; by exhibiting consternation that it wasn't yet done, and Rachael, well, just joyfully began exercising her innate capacity to develop the grammatical module, hence lending special support to the assumptions herein.

Special thanks to Irv Kessler, my favorite behavioral ideologue whose principles usually work in the right framework. And finally, a

tribute to Antonia Turman, who began as "the typist," a talented one at that, and grew into a supportive, special friend, who I've come to know as a gentle person, a scholar and an elegant lady.

VITA

October 4, 1948--Born, Toledo, Ohio

1970--B.A. (with honors), University of California, Riverside

1974--M.A., University of California, Los Angeles

1977--Distinguished TA Award, University of California, Los Angeles

1977-1978--Chancellor's Dissertation Year Fellowship

PUBLICATIONS

Klein, S. (1977) "A Base Analysis of the Floating Quantifier in French," presented at NELS VII and published in the proceedings.

Klein, S. (1977) "Linguistic Theory and a Residual Problem in the Acquisition of Syntax," MS, Department of Linguistics, University of California, Los Angeles.

Klein, S. (1978) "Predicting the Shape of Child Grammars: An Argument for the Role of Linguistic Theory in Language Acquisition," presented at WECOL and published in the proceedings.

ABSTRACT OF THE DISSERTATION

Syntactic Theory and the Developing Grammar:
Reestablishing the Relationship between Linguistic Theory
and Data from Language Acquisition

by

Sharon Michelle Klein
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 1982
Professor Stephen R. Anderson, Chair

This study brings together two sources of knowledge about the development of linguistic competence: linguistic theory and the observation of phenomena in child language. The central assumption is that the ontogenetic development of grammar is subject to the same principles required by a linguistic theory seeking to define the notion "possible grammar of a human language." The major goal is to determine what interaction of the data available to the hypothesis mechanism and the principles defining a theory of grammar will culminate in some intermediate grammar as well as how the same type of interaction allows the development of the adult grammar.

The theory of grammar fundamental here is that which has grown out of work directed toward constraining the functioning of transformational rules (Ross 1967; Schwartz 1972; Chomsky 1973, 1977, 1980, 1981;

Emonds 1976; and Baltin 1978). We assume the notion Universal Grammar (Chomsky 1977, 1980, 1981), that a number of the principles argued for in these works are part of Universal Grammar (UG), and that it is this set of principles that form a crucial part of the hypothesis mechanism.

The interaction of available data and these principles is shown to culminate in fragments of intermediate grammars accounting for a number of observable phenomena, such as the variable appearance of Subject-Auxiliary Inversion (SAI) in YES-NO and WH-questions in child language, and the analysis by children of Prt NP sequences as instances of PP. These fragments of intermediate grammars are, further, evaluated and justified on the basis of two criterial notions, Dissonance and Delearnability, which are argued to be necessary in the study of acquisition and a consequence of looking at acquisition in the framework outlined.

The framework is shown to allow us further to look at related acquisition problems in German and French. In the German case, the absence of SAI phenomena like that observed in English is provided a unified account, while in French, a prediction is made for development of the fragment of grammar accounting for certain aspects of the distribution of clitics.

Chapter 1

LINGUISTIC THEORY AND THE ACQUISITION PROBLEM: APPROACHES AND MISINTERPRETATIONS

1.0. Introductory Remarks

Maintaining that a central goal in the construction of a linguistic theory is to have the theory account for the ability of the human organism to acquire a language constitutes no revelation. Such a goal is stated quite explicitly in Chomsky (1964) where it is asserted that "an objective for linguistic theory is the precise specification of [an] abstract device . . . serving . . . as a model of language acquisition" (p. 61). The framework within which this goal is stated specifies the task of acquisition as the construction of a generative grammar on the basis of primary linguistic data. So stated, the goal in question becomes defining the mechanism controlling grammar construction.

This goal has been subject to interpretation in two general areas. Within the bounds of linguistic theory, attention has been focused on the attempt to constrain what are admissable as generative grammars. A major result of this attempt has been the determination of a set of principles constituting a Universal Grammar which serve to define a limited class of possible grammars. These principles would then represent the capacity of any human to construct for himself the grammar of

his language. A distinct interpretation of the goal is found in the area of developmental psycholinguistics where the determination of the acquisition capacity is claimed to be investigated more directly.

Initially, work in developmental psycholinguistics did embrace the central goal of linguistic theory and accepted the framework within which the goal was stated. In an often-cited work, McNeill (1970) states:

> Our theory of language acquisition will be that the theory of grammar and its universal constraints describe the internal structure of LAD [Language Acquisition Device] and, thus, of children. (p. 71)

Even prior to this quite precise statement of the goal, it was acknowledged that any account of the development of linguistic competence would require descriptive mechanisms distinct from what were available in general psychology. Bellugi and Brown (1964) maintained that "Somehow . . . every child processes the speech to which he is exposed so as to induce from it a latent structure." They illustrate this assertion by tracing the development of the noun phrase in child speech, and conclude that:

> It looks as if [the induction of latent structure] will put a serious strain on any learning theory thus far conceived in psychology. The very intricate simultaneous differentiation and integration that constitute the evolution of the noun phrase is more reminiscent of the biological development of an embryo than it is of a conditioned reflex. (p. 99)

Thus there seemed to be a good deal of enthusiasm for a framework that would make it possible to examine the development of language in light of what is determined specifically to be characteristic of language, rather than in light of some more general notions of cognitive development. Further, McNeill's mention of the LAD, a construct

introduced by Chomsky (1964) to illustrate the definition of a theory of grammar and its crucial relationship, as he saw it, to an explanation of language acquisition, makes it quite clear that this enthusiasm extended to the acceptance of the notion that such a line of investigation would require the postulation of nonlearned--innate--structures particular to language. In another paper, McNeill (1966) suggested that the early and universal appearance of constructions in the speech of children that must clearly be described as hierarchical supports the assumption that there are such innate structures:

> The essential feature of hierarchical structures--that they are hierarchical--is abstract and completely unmarked in overt speech, yet children always discover these structures. In order to account for this feat, a principle fundamentally different from inference is required. (p. 45)

Even while these early statements and reflections of the goals set for work in developmental psycholinguistics seem to have been in accord with those set for the construction of a linguistic theory, a subtle but crucial distinction is discernable. In linguistics, the specification of a theory of grammar (and hence the LAD) followed from an interaction of certain criteria set for the grammars of languages--that they be observationally and descriptively adequate--and the data themselves--roughly, native intuitions about grammaticality and related notions. Developmental psycholinguistics, on the other hand, recognized the validity of postulating a construct such as the LAD on linguistic grounds, but went about the tasks of specifying its content in a quite different way. The difference is best perceived by referring to the schema which illustrated the definition of LAD originally (Chomsky 1964), and comparing it to one we can construct for the task

defined in developmental psycholinguistics. The schemata appear as (1) and (2), respectively:

1. primary linguistic data → [LAD] → generative grammar
2. primary linguistic data → [?] → children's utterances

We can see fairly clearly that the specification of the mechanism in (2) will be quite distinct from that in (1). (2) most closely resembles a schema that would illustrate the general notion of grammar, but it does not reflect even this notion. The output of a grammar is a set of structural descriptions that reflect the information necessary to account for some speaker-hearer's interpretation of any arbitrary utterance, "to the highly non-trivial extent that understanding is determined by the structural description provided by the generative grammar" (Chomsky 1964, p. 6).

Such a schema as (2) characterizes the result of a research program with an initial goal of describing child speech. And this was the starting point:

> The first order of business would then seem to be to describe structurally the utterances children produce at various stages of development. One can then observe what parameters of the structure of the language the child seems capable of using to generate utterances. (Menyuk, Sentences Children Use, 1969, p. 9)

The obvious question to ask is why (2) should be the result of these starting points. Steps toward an answer to this question give us some insight as to precisely how, even given the apparent congruence with the goals set for linguistic research, developmental psycholinguistic research seems to have diverged from these goals. In an attempt to take these steps, I will, in the following sections, present brief

accounts of three approaches that crucially involve extensive descriptions of children's utterances and which, from these descriptions, make claims about the nature of children's competence and about the shape of (an) LAD.

1.1. The Interpretation of the Goals of the Theory by Early Studies

1.1.1. McNeill and Menyuk

The first approaches we will consider are McNeill (1966, 1970) and Menyuk (1969) to which we have already referred. Menyuk's work describes the development of syntax, largely in normal children, but with some discussion of abnormal language development. She points out problems associated with attempts to construct such descriptions: What the breadth of a corpus should be for a description to be assumed, how inappropriate descriptions might be inferred, and what the relationship of children's utterances might be to inferences made about children's linguistic competence, and the result of her work is a description from which such inferences are made.

In the course of the description, Menyuk traces the appearance of utterances (in a cross-section of 150 children from ages 3 to 7) which would reflect the development of phrase structure rules and subsequent development of transformations. To give, in her view, her descriptions greater reliability (validity also in her view) there are two experimental investigations (using the same sample of children) also reported. One was a repetition task in which children were asked to imitate sentences. Some of the sentences presented to them were

deviant, and the report of their performance in this test included which sentences they modified--changed the string in any way, and which ones they spontaneously corrected, as well as which ones were imitated with no changes. The second experimental task specifically asked children to correct deviant sentences. To the extent that the results in these experiments implied competence with respect to the same constructions that appeared in the spontaneous speech she described, Menyuk concluded that there was support for the descriptions.

It is important to note that Menyuk does not make an assumption that was made in some of the earliest and most influential works in developmental psycholinguistics. McNeill (1966, 1970) suggested that children acquire base rules before acquiring transformations. He claims in fact, "One can say that children begin speaking underlying structures directly." This claim was made within the framework of a grammatical model in which all semantic interpretation was done in the base, and hence the suggestion followed that the appearance of children's early speech to be largely semantic might be explained (1966, p. 51). (In the same model, of course, phonological rules interpreted surface structures. The apparent exception here was explained by the suggestion that children were producing "inner-speech" in the sense of Vygotsky 1962.) Compare the following:

> If children begin their productive linguistic careers with a competence limited to the base structure of sentences, it is difficult to see how it can be explained by any theory of language acquisition that restricts attention to what a child might obtain from the observable surface characteristics of parental speech. Such theories would have to predict the opposite course of development: First, surface

structure; then base structure. Most behaviorist theories have assumed this order, with notable lack of success; failure is inevitable when children produce only the base structure, and behaviorist theories produce only the surface structure of sentences. What is needed is either a child who commences acquisition with surface structure, or a theory that focuses on base structures. Since it is easier to change theories than children, the latter course has been followed here. (McNeill 1966, p. 52)

[The following refers to these sentence pairs]:

1. Child: Where Uncle Nat?
   Mother: (No reply)

2. C: Uncle Nat, Sylvia uh school
   M: Sylvia's working

3. C: Sylvia go to work?
   M: Yes

4. C: and Daddy goes uh work
   M: (No reply)

5. C: No goes uh work

It is apparent that rules are being applied to change the structure of sentences. . . . The question that has been raised is whether or not these should be considered transformational rules . . . if these utterances are non-transformational the child is formulating structural descriptions in his grammar for which he has no evidence. . . . Therefore, one cannot postulate that acquisition of syntactic rules can be accomplished or, indeed does take place, via a stimulus-response paradigm. On the other hand, exactly the same position can be maintained even if these utterances are produced with transformational rules. . . . Theoretically, it would seem logical to suppose that transformations are a part of the child's grammar in some form at the beginning stages of language acquisition since it is this aspect of the syntax which allows for the possibility of an infinite set of utterances and also allows for expressing different meanings using the same base structure rules (that is, 'Daddy goes to work,' versus 'Daddy doesn't go to work'). (Menyuk 1969, pp. 69-70)

We can see that McNeill's initial goal was to show that a stimulus-response theory of acquisition is inadequate because it would fail to

account for children's obvious ability to abstract away from what is "input" to the LAD. Menyuk (correctly) pointed out that assuming children "speak 'deep-structures'" first is not necessary to making such a claim. I would go further to point out that in fact McNeill has a problem if he does not assume the existence of the transformational component at a very early stage. Even if it were the case that children "spoke deep structures"; they are, as McNeill himself acknowledges, exposed to surface structures in their environment, and they would thus be abstracting away from these surface structures in order to speak deep structures. The mapping system between the two sets would have to be explained in some way. A logical proposal would be that the rudiments of a transformational system exist early on. This would not seem outrageous given the assumption that the framework for hypothesizing transformations is innate.

The failure to suggest the existence of such a system in McNeill is most likely traceable to two sources. The first is the assumption about what is actually universal and therefore innate. McNeill concentrated his attention on the universality of grammatical relations and the base rules that were supposed to express these. Transformations were not universal; not innate. The second is related to the first, perhaps in a way that makes it difficult to distinguish the two. This is the "feeling" about a model of grammar that was prevalent at the time and led to a number of interpretations of assumptions about its role in theories in psycholinguistics and language acquisition. The "feeling" was that the model was somehow dynamic; it could be applied directly as a model of production (and sometimes as one of

comprehension as well, hence of performance, generally), and that it was also in some sense directional.

This view is visible in Menyuk's work. It is quite possible that Menyuk did not propose transformations in the passage cited here because her view of the grammatical model prevented anything but a unidirectional view of a transformational grammar. Despite the objection she correctly raised, her description did proceed from base structures to the development of transformations. In fact we can see the development of a view of the role of the theory in the study of language acquisition here. Children produced utterances and the "hardware" of the theory was pressed into service to describe these utterances. It is the interpretation of these notions about a model of grammar and the role of the theory that led to some of the later criticism we will be considering.

A rigorous analysis of the data of child language may be a necessary step toward the development of a theory of acquisition, but it is by no means sufficient to derive such a theory. There are at least two reasons for making such an assertion. Even if an extensive description of children's speech (buttressed, albeit with certain types of experimental evidence) could give us a clear picture of their grammatical competence, this picture of competence is not isomorphic with any theory of grammar, and hence, any conception of the LAD. Nor is such a picture even necessarily reflective of a theory. The second reason, therefore, for the insufficiency of a description as the crucial basis for a theory is that for a description to give us any insight into the substance of the LAD it must have justification

external to the data for which it is a description. When descriptions, as we have seen, are motivated exclusively by the data they are describing, the picture of the LAD which we get is either hazy, which is largely the case in the Menyuk study, or is grossly inconsistent with the assumptions underlying the linguistic theory, which ends up providing the descriptive hardware but little else; the case primarily in McNeill's work, but also visible in Menyuk.

1.1.2. Smith and The Acquisition of Phonology

A third study underlines the two problems we have raised for attempts to derive a theory of grammar construction (a picture of LAD) from a rigorous description of child speech because it succumbs to the kind of conclusions they invite in a fairly spectacular way. Hence, we will see quite clearly the divergence of developmental psycholinguistic studies, characterized by the three we will have considered here, from the original goals stated for the determination of a linguistic theory.

In The Acquisition of Phonology (1973), Neilson Smith proposed a detailed description of his child's phonological system. Outlining and rejecting an account of the child's phonology as a "self-contained system," Smith proceeded then to detail a description of which a major feature was a set of phonological rules ("realization rules") operating on adult surface forms to derive 'child' phonemic forms which were then subject to certain phonetic rules. These rules, then, provided an account of the child (fairly gross) deviations from adult pronunciations. First, it is not clear what Smith meant by adult surface forms.

He referred to them at times as "surface phonemic forms" or as "adult phonemic forms," taking pains to distinguish them from "underlying forms," in the sense of Chomsky and Halle (1968) (cf. Smith, pp. 178 ff.). This problem is important because Smith claims that "the child's competence is a close reflection of the adult form he hears and . . . his deviant output is the result of the operation of a set of psychologically valid realization rules" (p. 133). Further, the realization rules are claimed not merely to reflect performance constraints, but to entail other features of the child's competence as well.

The picture of the child's phonological competence is more obscured. In Smith's work as well as in McNeill's and Menyuk's, the 'direction' of the grammar becomes a problem. As Smith himself acknowledged in a note on the way his system might relate to a psychological model of acquisition, the realization rules (and I suppose the phonetic rules, though he did not specify them explicitly) "could map a surface (phonemic) adult form such as /skwi:z/'squeeze' onto a child utterance [ki:b], but not vice versa" (p. 183). Thus a detailed description of a child's phonological output does not give us a very good picture of the child's phonological competence. Not only is the picture hazy, it is also inconsistent with a central requirement for such a picture in the linguistic theory whose hardware the description uses: no grammar produces utterances (or "surface forms") from some other more abstract forms.

The second problem is visible in this work as well. From the view of competence provided by the description we briefly sketched, Smith makes proposals about the content of the LAD. And, as we

claimed is the outcome in such cases, the content of his conclusions bears little resemblance to an LAD consistent with the theory of grammar that defined it.

> The realization rules correspond to the hypotheses the child makes about the language he is learning, and changes in the realization rules through time correspond to fresh hypotheses made by him. . . . I suggest that they [the hypotheses] are limited by the set of distinctive features and by pressure from four, hierarchically arranged, universal tendencies. These tendencies or constraints are:
>
> (i) Vowel and consonant harmonization
> (ii) Cluster reduction leading to a CVCV . . . canonical form
> (iii) Systematic simplification
> (iv) Grammatical simplification
>
> That is, one can view these constraints as being part of a universal template which the child has to escape from in order to learn his language. Changes in the realization rules then represent the idiosyncratically variable attempts of a given child to overcome the universal characteristics of language in his progress toward learning his own specific language. (Smith 1973, p. 206)

The conclusion here is quite distinct from a notion of the substance of an LAD as schematized in (1) that would have largely the same content as a theory of grammar as defined within the bounds of linguistic theory. If one concludes that the acquisition of language involves (crucially) escaping the boundaries of certain universal constraints, it seems bizarre to suggest that these are the same constraints which limit the class of grammars allowed to be tested as descriptions of language.

The outcome of what at first seemed minor distinctions in linguistic theory and in developmental psycholinguistic work that embraced the goals of linguistic theory is fairly stunning, then. I have tried to point out what the differences were, to show that they

were, although subtle, fairly fundamental, and that the consequences--in terms of conclusions drawn in the field of psycholinguistics about language acquisition--were a profound disappointment. A good deal of criticism of linguistic theory and its role in the investigation of language acquisition ensued; no doubt in response to the disappointment. We will now focus on the content of some of that criticism, in an effort to determine the extent to which it is actually linguistic theory that is being criticized or rather the wrong assumptions about linguistic theory that characterized developmental psycholinguistic studies such as the three we have briefly examined.

There is another source for the criticism that is relevant here, and that is misinterpretation of certain of the assumptions. This is, I believe, particularly true of criticisms that we can trace to studies such as McNeil's (1970), such as that of Bowerman, which we will consider below. The misinterpretation runs along the following lines. Very roughly put, McNeill's claims are that the more "basic" and therefore universal features of language are innate and thus appear first. Specifically he maintained that grammatical relations are universal and sentences reflecting them directly--base generated strings, the rules of the base being universal--appeared first. It is more the details of this claim that invalidate it, rather than the principle motivating it: that there is a rich innate component to the language acquisition mechanism. What we consider mistakes about the content of the innate component and its (direct) observability in this account have as much to do with the history of the development of linguistic theory as they do with what we claimed above is an incorrect

assumption--that a model of grammar is a dynamic model of production or of acquisition, for that matter.

The serious error, in my view, is one that was not made by McNeill. Rather, it was made by his critics, who were also involved with arguing against the necessity of connecting the construction of a theory of language acquisition with a theory of grammar; in particular Generative Transformational Grammar. The error lies in the omission of one step in the logic of his claim. This omission results in the claim that structures reflecting the P-S rules are more "basic" and are therefore more likely to appear earlier. The significant missing step is of course the claim about universality. This error results, I believe, from a combination of the interpretation of McNeill, involving the missing step, and the view projected in the theory of acquisition derived from the Derivational Theory of Complexity. In this view a structure that was claimed to involve $\underline{n}$ transformational rules was acquired earlier than one that required $\underline{n} + \underline{1}$ transformations (Brown and Hanlon 1970). The results of this general misinterpretation are visible in a number of the criticisms we will consider in the following sections.

1.2. Criticisms of the Theory: A Reaction to the Early Studies

The criticisms were of two types, generally. On the one hand, certain critical points of view were outlined to lead directly to counterproposals for descriptions of child language. The second type had no specific counterproposal in mind. Outlining what it viewed as

the failure of a particular linguistic theory as an "explanatory tool" in the area of developmental psycholinguistics, it became an assertion that linguistic theory should be abandoned as such a tool in this area.

1.2.1. What is Innate? Misinterpretations of the Claims of Linguistic Theory

A thread common to conclusions sustaining the criticism in both types is the identification of a particular model of transformational grammar with a theory of grammars, and further, the confusion of the substance of a theory of transformational grammar (at some point in its development) with the program outlined to determine the substance. While the beginnings of such confusion were visible in the work we considered in section 1.1 above, the statements revealing it became quite explicit in subsequent work.

Bowerman (1973) saw the role of a particular theoretical framework to be one of determining what kind of grammars one would write for child language. She claimed that the assumptions underlying transformational generative grammar led investigators to posit (3) as the set of base rules characterizing children's early three word (Noun Verb Noun) utterances.

(3) a. $S \rightarrow NP\ VP$
b. $VP \rightarrow V\ NP$
c. $NP \rightarrow N$

She cited McNeill in support of this claim:

> According to McNeill (1971), this hierarchical organization of sentence constituents results automatically from the child's application to sentences of his innate knowledge of the basic grammatical relations. (Bowerman, p. 178)

She argued further that because transformational theory held grammatical relations to be stated at deep structure, and because deep

structures are "inaccessible" to children, these relations are unlearnable. Hence, her argument continued, within the framework of transformational grammar, the expression of these relations must be inferred to be innate. Her criticism of the framework then focused on the assertion that the VP constituent, which (3) assumes, does not have linguistic justification in the early speech of children. Linguistic justification she defined in terms of a footnote in Chomsky (1965) (p. 197, note 7), where he suggested that the claim of constituent structure for a coordinated string is supported by showing that such a structure is "required for some grammatical analysis, that the [phrase] must receive a semantic interpretation, that they define a phonetic contour, that there are perceptual grounds for the analysis, or something of this sort." This note is in reference to an outline of what might be involved in the study of performance--in particular, what parameters of superficial structure might set limits on acceptability--thus limiting performance in some way. Chomsky cited as one "plausible" observation (4iii, p. 13), "multiple branching constructions are optimal in acceptability." He suggested that this high acceptability might be accounted for by some representation of how much "computation" an analysis of a sentence would require and that it was reasonable to suggest that the ratio of phrase node (e.g., NP, VP) to terminal nodes (N, V) would be such a representation.

It seems that this criticism levelled against the role of linguistic theory in language acquisition of Bowerman (1973) was based on at least two assumptions about transformational theory that are not in fact part of that theory. The first assumption seems to be that rule

schemata such as (3) are innate, or, more weakly, that a constituent VP is innate. Neither the strong or weak form of this assumption is part of the theory, even though it does seem to be a fundamental assumption in acquisition work. It is most plausibly this assumption that led McNeill to the notion referred to in 1.1.1 that the model of grammar could be interpreted as a model of production in some sense. It certainly can be seen as underlying the claim that children utter deep structures first and then acquire transformations. This claim is still very much a part of acquisition research, and continues to surface in discussions of the explanatory role of linguistic theory. Consider the following passage from deVilliers and deVilliers (1978).

> It would be particularly nice for the theory if the child first produced sentences rather like the proposed deep structure, then acquired the transformations to produce a closer and closer approximation to the surface form that adults use. (Of course the deep structures proposed by linguists are not given phonological form until the very last step in the derivation.) (p. 102)[1]

The second assumption is really a corollary of the first, but it reveals more clearly the confusion of an interpretation of a particular model of grammar proposed within the theory, with the theory itself. The second assumption is that actual base rules such as those in (3) are universal because they express grammatical relations and these are universal. Therefore such rules are part of the biological specification that the child brings to acquisition. It does not seem unreasonable to question the correctness of such an assumption, even without testing the predictions it would make for the shape of child grammars. A rule schema such as (3) can quite readily be shown to be inadequate to describe languages with word orders distinct from

English. And, of course, children acquire these as well as English. Further, it does not seem unreasonable to suggest that the constituent status of the VN sequences in VP is inferred by the child's mechanism for grammar construction. There is nothing in the theory (or the model in question, that is, the Aspects model, for that matter) that rules out the possibility of a language with no VP constituent.

There is no feeling in the statement just cited from Chomsky that such constituents as VP must be heavily justified or, on the other hand, that they are the only possible ways of describing N V N sequences. Rather, the suggestion we may make from the statement is that, all things being equal, there is a predisposition toward the construction of grammars that involve such constituent structure as (3a), while descriptions such as S → NP V NP are not ruled out.

In fact, the universal status of the base rules was defined in Chomsky (1965) in the following way:

> To a large extent, the rules of the base may be universal, and thus not, strictly speaking, part of particular grammars; or it may be that, although free in part, the choice of base rules is constrained by a universal condition on the grammatical functions that are defined. (Aspects, p. 141)

This definition seems to be a proposal for further research concerning the parameters within which base rules may vary rather than an assertion basic to the theory. In fact, the recent proposals for the shape of base rules within the framework of $\bar{X}$ syntax (cf. Jackendoff 1977, Selkirk 1977) are more consistent with the second suggestion in this statement than with the first. There have been substantive changes in the 'model'; the determination of its substance was clearly set as an empirical problem. Abandoning the theory on the basis of inadequacies

of the model--be they actual inadequacies (which was clearly the case in some instances) or be they inferred because of particular interpretation of what was included in the model--does not seem justified by itself.

As mentioned above, however, the criticisms of which Bowerman's serves as an example introduced proposals for alternative ways of describing child language. Braine (1976) tries to make a stronger argument against linguistic theory based on the same assumptions as Bowerman's. He writes that while Bowerman based her argument on the lack of justification for the VP constituent, he could cite evidence against the postulation of such a constituent. His arguments involve two-word utterances while Bowerman's concerned three-word (N-V-N) utterances, as well as ones consisting of two words (e.g., N V; V N). Briefly, he argues that a rule like S → NP VP for "actor-action" sentences makes it impossible to account for the two-word orders present in noun-"locative" utterances (N-LOC and LOC-N). Of course, we can point out again that the theory does not limit the child's grammars to a rule of the type S → NP VP. Further, great care should be taken in the inference of grammars from child utterances. LOC-N sequences may be instances of PP, not instances of S at all, and thus will be irrelevant to any arguments of the type Braine is making.

In any case, both Bowerman and Braine support alternatives for describing child language that are of the same type, although different in degree. Bowerman adopts a system close to the case grammar proposed by Fillmore (1968), while Braine rejects this approach for its "poor fit" with the corpora. He proposes

> that children are learning, seriatum, a number of formulae of limited scope, each formula being a rule that maps elements of a semantic representation into positions in the surface structure. The formulae are limited in scope in that each is concerned with a specific, often quite narrow, range of relational conceptual content. (Braine 1976, p. 69)[2]

Nonetheless, both see the question of whether to use descriptions involving syntactic notions or whether to focus on what they see as more primary semantic ones as a central issue. In a commentary on Braine (1976), Bowerman writes:

> There has been much recent debate over the issue of whether children base their initial rules for combining and ordering words purely on semantic notions like 'agent' and 'processor,' or acquire instead an early understanding of non-semantic (syntactic) relationships such as those holding between subject and predicate. . . . The ultimate resolution of the issue (assuming one is achieved) will contribute considerably to our understanding of the level of abstraction at which children's initial strategies for processing language operate. (commentary in Braine, p. 100)

Bowerman's assessment reveals yet another theme running through most of the counterproposals to descriptions that are seen as derived from what is interpreted as linguistic theory. That is the notion that basic to the understanding of how children acquire language is an understanding of what children's strategies for processing language are.

The introduction of the notion of processing (whatever its definition turned out to be), broadens the scope of what factors were seen as central to the development of language. Brown (1972) proposes (in terms reminiscent of Piaget),

> that the first sentences express the construction of reality which is the terminal achievement of sensory-motor intelligence. What has been acquired on the plane of motor intelligence (the permanence of form and substance of immediate

> objects) and the structure of immediate space and time does not need to be formed all over again on the plane of representation. Representation starts with just those meanings that are most available to it, propositions about action schemas involving agents and objects, assertions of non-existence, recurrence, location, and so on. (p. 200)

Further, Bloom (One Word at a Time, 1973) maintains that a "rich interpretation" (an interpretation that allows for the expression of a fairly complex situation by the child in one word utterances) does not require us to attribute to the child "the linguistic knowledge for talking about such relationships" (p. 137). She proposes that "children develop certain conceptual representations of regularly recurring experiences, and then learn whatever words conveniently code such conceptual notions" (p. 113).

Thus from the assertion that some sorts of semantic notions, as opposed to linguistic categories, are fundamental to acquisition, the question of the source of these notions has led these investigators to propose that they are, in some sense, primitives of cognition, and, that the child expresses them, in some way. Slobin (1978) takes a position that is slightly more moderate. He uses a "waiting room" metaphor, where the entry key is an underlying conceptual notion and the exit key is the appropriate linguistic form to express the notion. As well, the child has in tow some sort of semantic notion (a key to the next room?) and the morphosyntactic information necessary to define the form (pp. 31-32).

A crucial observation to make here is that in none of these counterproposals is there an explicit denial that some sort of generative transformational grammar is the endpoint of acquisition for adult competence, and therefore is the endpoint of acquisition (but,

cf. note 2, p. 88). The problem that has no solution in any of these approaches, therefore (where cognition triggers linguistic development) is how it turns out that children's systems develop into precisely the kinds of grammars that are justified as descriptions of adult competence.

It is not difficult to trace the sources of this discontinuity. The concern of research in the area of language acquisition has shifted. The studies we cited in Section 1.1 held that a first step in research was a rigorous description of child utterances. The studies we have just discussed had no explicit goal beyond this description. Grammatical models were evaluated on the basis of how well they adapted themselves to describe the things children were saying. This evaluation framework is particularly apparent in the work of Bowerman. Here too, we can see the development of the argument against linguistic theory in these studies.

We have already made a brief reference to Bowerman's justification for judging models on these grounds in the discussion of the phrase structure rule schema in (3) (cf. p. 14). Generalizing from that discussion we can characterize the justification in this way. The assumption basic to linguistic theory was taken to be the notion of innateness. A number of researchers (among them, McNeill at the time) interpreted this notion as a claim that base structures were actually innate. We have seen the line of reasoning; deep structures (as the term was defined first--untransformed structures) are abstract to the extent that there is no way to get to them directly from the source structures. From the speech children hear, deep structures are

unlearnable, hence, innate. This argument led to the prediction that children's early utterances should resemble the output of phrase structure rules.

The line of argumentation had no real support from any of the assumptions underlying linguistic theory. But we can see how the assumptions were interpreted to support the argument. The connection between unlearnable and innate is the key. In fact the interpretation of 'unlearnable' that seems correct to impute to linguistic theory is 'not describable in terms of any general learning theory that might exist.' As Chomsky states (1965, pp. 57 ff.):

> We have a certain amount of evidence about the character of generative grammars that must be the 'output' of an acquisition model for language. This evidence shows clearly . . . that knowledge of grammatical structure cannot arise by application of step-by-step inductive operations (segmentation, classification, substitution procedures, filling slots in frames, association, etc.) of any sort that have been developed within linguistics, psychology, or philosophy. (p. 57)

A generative grammar (given that this is the output of acquisition) is not therefore unlearnable in the absolute sense, but rather requires a very rich set of "guideposts" to ensure that it, rather than some other of the indefinitely many types of descriptions imaginable,is the result of acquisition--the result of learning. The theory claims then, that learning is restricted in quite specific ways. It is restricted by a set of principles and definitions (what we referred to informally as the "guideposts") that place limits on the shape of hypotheses in the course of the "learning" of the grammar of a particular language. These principles and definitions are claimed to be part of an "initial, innate structure" and are not learned. The

problem set for linguistic theory is

> that of developing a hypothesis about initial structure that is sufficiently rich to account for acquisition of language, yet not so rich as to be inconsistent with the known diversity of language. (Chomsky 1965, p. 53)

The position is that the characterization of initial structure is not learnable, in the absolute sense; there is no direct relation between it and the "environmental experience" relevant to acquisition--the language that the child hears.

How the confusion between the theory of grammars and a model of grammar (a particular theory of a language) is reflected in the criticisms of the theory and in the work which elicited them should now be clearer. The criticisms considered it a failure of the theory when children were not observed to utter exclusively what were supposed to be base structures. From another point of view the argument was made that even one observation that most of a child's utterances "<u>could</u> be generated by the base structure rules of a transformational grammar without the intervention of transformational rules" (Bowerman 1973, p. 175) did not "constitute evidence that children have innate linguistic knowledge corresponding to the abstract and unobservable base structure representations of sentences" (ibid). These criticisms and the disillusionment with a theory they reflect come precisely from the inappropriate identification of the theory with the grammar in the specification of what was claimed to be learned and what was claimed to be innate. The predictions were doomed to be inaccurate because they were based on ill-made assumptions. What is surprising is not that observations bore out the inadequacies of the predictions, but

that people saw this outcome as the result of a failure of the theory as an explanatory tool. But it was so interpreted. Thus work turned explicitly to the description of child language. The turn to semantic notions as the basis of a descriptive mechanism was therefore based on how well such a description fit the child data, not on the role such notions might play as a mechanism for describing the course of language acquisition.

1.2.2. What is the Theory a Model of? The Fodor, Bever, and Garrett Criticism and Its Source

In 1.2, the criticism of linguistic theory as a potentially explanatory tool in the study of language acquisition was characterized as generally being of two types. The foregoing discussion traced how proposals for alternative approaches to the acquisition problem, and in fact, goals distinct from those stated for the theory grew out of such criticism. A criticism of the second general type appears in The Psychology of Language (Fodor, Bever, Garrett 1974). There is no unique or consistent proposal for an alternative to the principles of the theory (or to the goals), but an argument is made against the theoretical framework. As I have suggested is characteristic of both types of criticism, much of the argument here appears to be the result of identifying certain aspects of a particular model of the theory with the program of research outlined by the goals of the theory.

Although reference has been made to this confusion, I believe the problem needs to be stated more explicitly. To begin, consider 4(a)-(c), from Aspects of the Theory of Syntax.

(4) a. An acquisition model

(i) a technique for representing input signals.

(ii) a way of representing structural information about these signals.

(iii) some initial delimitation of a class of possibile hypotheses about language structure.

(iv) a method for determining what each hypothesis implies with respect to each sentence.

(v) a method for selecting one of the (presumably, infinitely many) hypotheses that are allowed by (iii) and are compatible with the given primary linguistic data.

b. A theory of linguistic structure

(i) a universal phonetic theory that defines the notion "possible sentence."

(ii) a definition of "structural description."

(iii) a definition of "generative grammar."

(iv) a method for determining the structural description of a sentence, given a grammar.

(v) a way of evaluating alternative proposed grammars.

c. (i) an enumeration of the class S_1, S_2, . . . of possible sentences.

(ii) an enumeration of the class SD_1, SD_2, . . . of possible structural descriptions.

(iii) an enumeration of the class G_1, G_2, . . . of possible generative grammars.

(iv) specification of a function f such that $SD_{f(i,j)}$ is the structural description assigned to sentence S_i by Grammar G_j for arbitrary i,j.

(v) a specification of a function m such that m(i) is an integer associated with the grammar G_i as its value (with, let us say, lower value indicated by higher number). (Chomsky, Aspects, 1965, pp. 30-31)

With 4(a)-(b), an explicit parallel is drawn between what must be included in an acquisition model, and what "a theory of linguistic

structure that aims for explanatory adequacy must contain" (Chomsky 1965). Four (c) represents a formal specification of the notions in 4(b). (4) hence, outlines requirements for a theory. In a way, (4) also outlines a program for determining the content of the theory. It is the various interpretations of the statements in (4), distinct from what we may infer is their intent, that underlie both the criticisms of the role of (any) linguistic theory in the study of language acquisition and most of the studies themselves, as we have just seen.

Fodor, Bever and Garrett interpret the statements in (4) as outlining somehow a process model of acquisition, rather than as setting out the requirements for determining the content of an acquisition model, which is not claimed to represent any actual learning procedure. Therein lies the confusion that seems to motivate the criticism. They claim that the model they see (4) as outlining can be thought to have the same characteristics as an analysis-by-synthesis model of speech recognition of which a central feature is a grammar which generates a string that is compared for match with some input string. In the course of acquisition some set of grammars would be scanned to see which one of them could generate the same strings as the input one. This interpretation is the source of the criticism from which the following is taken:

> Roughly speaking, Chomsky's model uses the input data only in the most inefficient manner. In effect, the only question that the child is allowed to ask about his corpus is whether it matches the predictions of some candidate grammar <u>at the phonetic level</u>. Though there is no reason for denying that distributional facts, semantic facts, facts of discourse structure, etc., may play a role in determining the child's developing picture of the grammar of his lan-

> guage, it is difficult to see how Chomsky's model could accommodate this possibility.
>
> It should be noted that this difficulty is inherent not in the notion that the child has innate information about the language he has to learn, but rather in the suggestion that the information is organized in the form of a recursive enumeration of the possible grammars. For example, these problems vanish if it is assumed that the child's innate information is organized as a set of "discovery procedures" for operating on a corpus to construct candidate grammars. (Fodor, Bever, Garrett, The Psychology of Language, pp. 475-476)

There is also a hint that Fodor, Bever and Garrett see in the outline a particular model of grammar. When they make claims about the child being limited to asking questions at the phonetic level, they must attribute some real model status to the way b(i) is stated, when it is, in fact, a proposal that a certain class of information should be represented in some way. One can propose a model of grammar in which sentences have a level of phonetic representation. In fact one must, of course, if particular generalizations not only about phonology, but about the interaction of syntax and phonology are to be stated. That particular "level of representation" will make no claims about what kind of information may play a central role in the assignment of structural descriptions to sentences, except in a very limited way. It will provide information of a certain form for the determination of structural descriptions. Whether the information is crucial, or of any use at some other "level" is a separate question, to be determined as the grammar is defined. To illustrate such a point, albeit in a slightly different direction, one can cite Bresnan's (1971) arguments that certain phonological processes, in particular, stress assignment, must have access to more syntactic information than

was previously claimed to be required.

Generally, while the substance of "structural descriptions" may change, the program set out in (4) does not limit these changes, it merely sets up the framework in which it may be determined whether changes are required. For example, the content of structural descriptions assigned to sentences has changed in fairly fundamental ways since Aspects, given the radical changes in the way the interpretive component relates to the syntax in the grammar. However, these changes do not affect the statements of (ii), (iii), or (iv) because these statements do not outline the substance of any model, but outline the requirements that any model must meet. Thus, a proposal for a certain technique for representing information, where it is the requirement for the information that characterizes the theory and not a specification of how the information is to be represented, can be abandoned in favor of an alternative without the abandonment of the theory.

More directly, one can think of possible interpretations for (b)i, given (a)i and (c)i that do not require the conclusion that a child is limited to making all inferences about the grammar he is acquiring at the phonetic level. For example, sentence types may be defined by unique intonation contours--reflected in the phonetic representation of strings. Once a sentence type is defined, facts about its internal structure, about its meaning, and about its status in the language given a particular grammar (i.e., is the sentence generable?) can be determined. In fact such an imaginable interpretation turns out to be consistent with what has actually been observed

in child language development. Children's prelexical utterances have been observed to include intonation contours of sentence types--declarative and interrogative, precisely. This behavior comes before the time that children enter the stage I level, referred to as the one-word stage (or holophrastic stage by those who maintain that these single words are sentences, but cf. Bloom [1973]). Although great care must always be taken before inferences are made on the basis of behavior, when a systematic set of behaviors corresponds[3] to what one would expect as the output of a mechanism that must satisfy a requirement defined as 4(a)i, (b)i or (c)i, the behaviors may be interpreted as supporting the hypothesization of such a mechanism, in other words, consistent with the Aspects' conception.

1.2.3. Denying the Validity of the Goals of Linguistic Theory: The Critique in Transformational Grammar as a Theory of Language Acquisition

We will consider yet one other criticism. This is the work of Derwing (1973), Transformational Grammar as a Theory of Language Acquisition. Derwing claims to challenge the assumptions underlying linguistic theory. Related as they are, both of the notions running through the language acquisition literature we have already examined converge in this work. These are, first, the notion of what is claimed to be innate in linguistic theory, and second (following from the first), what kind of mechanism is required to "learn" language. Derwing denies that there is any set of inherent, unlearned principles specific to language that can be thought of as part of a learning mechanism for language. The following reflects the essence of his argument.

> Suppose . . . we argue, with Chomsky, that transformational-generative grammars (and the associated Universal conventions that go with them) are incapable of being 'learned' in any presently understood sense of the term. What conclusions are to be drawn from this? To my mind the answer is straightforward: If grammars of the Chomskyan sort cannot be learned by any means presently known, such grammars simply cannot be accepted as plausible or realistic models of any actual psychological entity or process. . . . To assume instead, as Chomsky has done, that the child must come to the learning situation fully equipped with a priori knowledge of some specific set of innate 'linguistic universals' which 'makes possible' the acquisition of such 'a rich and highly specific system' as he proposes for grammar is counterproductive, since this amounts to defining the problem away.
>
> I conclude, therefore, that despite the number and range of claims advanced for linguistic descriptions cast within the abstract transformational-generative framework, these descriptions are inappropriate to the output of the process of language acquisition, and that serious metatheoretical and methodological difficulties must be inherent in the whole Chomskyan approach to the study of language. (pp. 69-70)

In sum, the argument runs as follows. If transformational grammars are too abstract and complex to be learned through the interaction of some general set of inference principles and the linguistic data in a child's environment, then transformational grammar cannot be argued to be an appropriate descriptive device for the knowledge that comprises human linguistic competence because there is no language specific learning mechanism to learn such a device.

The mechanism of learning, should, Derwing holds, be characterized by what he calls a "process model," comprised of some general learning algorithm, or some data analysis mechanism. This view he contrasts with a "content" model, which is described as including "inate universals plus evaluation measures." Derwing attributes the terms "process" and "content" and the contrast between the two views

of a language learning mechanism he holds they reflect to Slobin, who first made use of them.

> Slobin himself favors what he calls a 'process' approach, in terms of which 'the child is born not with a set of linguistic categories but with some sort of process mechanism--a set of procedures and inference rules, if you will--that he uses to process linguistic data.' Under such an interpretation as this, then any linguistic universals would be 'the result of an innate cognitive competence rather than the content of such a competence.' (p. 88)

(The sequence inside single quotation marks are Derwing's citation of Slobin 1966; Derwing, p. 54.)

Derwing attributes to Slobin an interpretation of the contrast between "content" and "process" that is nowhere evident in Slobin's definition of the two. Slobin was not denying the existence of a language specific acquisition device, he was challenging one view of its substance: that of McNeill in the same volume. McNeill, recall, maintained that the device includes precise details of the hierarchical arrangement of syntactic categories, in addition to information about the categories themselves, i.e., phrase structure rules. What is more significant is that Slobin nowhere asserts that the "set of procedures and inferences rules," the "process model," is not language specific.

However, it is precisely that assertion that Derwing attributes to Slobin. Derwing rejects the notion that it is fruitful to hypothesize a language specific acquisition device, and suggests that a "process" view of such a device is not language specific. There are two major problems with this argument of Derwing's. First, as we have seen, its premises are founded on an incorrect interpretation of

Slobin's contrast between "content" and "process."[4] Slobin's contrast derives from his criticism of a quite specific realization of a "content" view: McNeill's. Second, it interprets this criticism as support for a non-language specific acquisition device, which the criticism is not. Derwing maintains that the incorrectness of McNeill's hypothesis (pointed out by Slobin) throws (serious) doubt onto the notion of an acquisition device such as that outlined in (4). Nothing in the fundamental nature of the theory, however (as we have seen), invited McNeill to propose the hypothesis about the innateness of categories and phrase structure rules that he did.

Even ignoring the added complication of the misinterpretation of Slobin by Derwing, the pattern of Derwing's argument becomes clear. His target is the theory of grammar and the framework for outlining how its content is to be determined: the essentials of (4). Derwing denies that there is any innate language specific acquisition mechanism, related to such a theory, which a child "uses" to determine the grammar of his/her language. He claims:

> The inherent advantages of the 'process' or 'empiricist' position are thus twofold: (1) this is the position which makes the fewer assumptions about the child's innate capacity for language and hence is to be preferred on conceptual grounds, all other things being equal; . . .; and (2) it is only this less complex and less restricted 'learning algorithm' approach which offers any hope of actually explaining language acquisition as a special case of some more general theory of human learning. (p. 55)

His arguments to support this claim, against the theory, are based on weaknesses that he points out in proposals which he claims are either crucial projections of the theory or are its (only) logical extensions. The essential flaws in Derwing's arguments lie in what he asserts are

"projections" on the one hand, and what he sees as the relationship between the theoretical framework and proposals that may, or may not in fact, be projections of it.

To begin with, Derwing fails to distinguish between (a priori) assumptions about the content of a language acquisition device and the construction of hypotheses about its substance within a theoretical framework (such as (4), for example). This particular confusion surfaces in his assertion that the following note from Aspects (Chomsky 1965) involves a contradiction (Derwing, p. 73).

> There are no grounds for any specific assumptions, empiricist or otherwise, about the internal structure of this device. Continuing with no preconceptions, we would naturally turn to the study of uniformities in the output (formal and substantive universals), which we then must attribute to the structure of the device (or, if this can be shown, to uniformities in the input, this rarely being a serious one in the cases that are of interest). This in effect, has been the rationalist approach, and it is difficult to see what alternative there can be to it if dogmatic presuppositions as to the nature of mental processes are eliminated. (Aspects, p. 207, n. 33) [Derwing cites parts of this passage (Derwing 1973, p. 73)]

It is only the blurring of the distinction which allows Derwing to see this as a contradictory statement.

Derwing also fails to distinguish between hypotheses about the shape of grammars that are projections of the theory and those that are not. For example, Derwing aims the thrust of his argument against Transformational Generative Grammar at its "unconstrained" nature (at any level other than possibly the level of transformations).

> There are no specific constraints on the notion of 'possible lexical representation' in generative phonology and on the notion 'possible deep structure' in generative syntax. Constraints on rules are meaningless so long as corresponding constraints are not also imposed on underlying representa-

> tions; the whole linguistic system must be constrained, not just that part that one happens to be particularly interested in. (p. 168)

The exhortation that the system must include constraints at all levels is an important and valid one. But it is not one that has ever been ignored in the development of the theory. Whether or not one accepts the substance of certain proposals (Jackendoff 1977; Selkirk 1977; Hornstein 1977), the system of $\bar{X}$ Syntax provides a fairly tight framework for establishing what may be a possible phrase structure rule. Thus, there is a very definite mechanism for constraining the notion 'possible deep structure.' Further, at other points in the grammar, attempts at determining the nature of the constraints that must be imposed on the "entire system" are quite serious. Constraints on (transformational) rules are broken down into constraints on the shape of the rules (what may or must be included in a structural description), on the manner in which they may apply (in obedience of a cycle, or not, for example), and on the shape of their output (structure preservingness, notions of 'landing sites,' etc.). As well, it is turning out to be the case that there are limits on what may be the input to interpretive rules in a grammar: that is what may qualify as an interpretable structure. And there is evidence that these constraints are distinct (at least in part) from those either on underlying forms or on transformational rules.[5]

There are two points to be made with reference to Derwing's statements about constraints. First, as should be clear from (4), the goals for linguistic theory were never defined in terms of a unique level, and attempts to constrain the theory were not confined to a

single level either. Thus, it does not make sense to speak of any constraint on the "whole system," if we take Derwing literally. Second, the goal of constraining the definition of possible grammar has always been justified from two points of view; that internal to the mechanism of the theory, and that directly related to acquisition: a system that allows an extremely wide range of grammars would fail to account for the (relatively) short time it takes for a child to achieve (nearly) adult competence and the (relative) "ease" with which she attains this competence. Thus there is nothing in the definition of the structure of the theory or in the statement of its goals that suggests it falls short of Derwing's demands for constraints. On the contrary, it is precisely such demands that are explicitly stated as goals.

However, what allows Derwing to find fault with the theory along these lines is his presumption that Generative Semantics is a well-formed descriptive hypothesis within the framework of the theory of grammar defined in Aspects; that it is a natural projection of the theory, given the Katz-Postal hypothesis. On the basis of this assumption, Derwing dismisses transformations as merely:

> devices which the linguist employs to derive diverse syntactic structures from the same underlying semantic (or semantically interpretable) 'deep' representation and (ever since Katz and Postal's book) without changing their meaning. (p. 162)

Neither such a definition of transformation, nor the position taken in Generative Semantics--that transformations derive surface structures from semantic representations--are logical consequences of linguistic theory, even one which includes the Katz-Postal hypothesis. They are

conclusions that can be drawn only from a particular interpretation of counterexamples to the hypothesis and a reformulation, based on this interpretation, of the hypothesis as an axiom of the theory.

Derwing's charges about the unconstrained nature of Generative Semantics, and of the trivialization of the notion 'transformational rule' in this model are not invalid. However, his charge that the inadequacy of this model reflects an essential failure of the theory is clearly invalid. Not only is Generative Semantics not a natural projection of the theory, it was in fact rejected, among other reasons, on the grounds that it failed to adhere to the very goals of constraining what is a well formed description (a grammatical model) (cf. Jackendoff 1973; Chomsky 1972).

Derwing compounds this failure to distinguish between hypotheses that are actual projections of the theory and those that are not. He blurs as well the distinction between goals set for a theory of grammar and goals set for a grammar. Derwing correctly states that the goals of a description include providing an account of certain linguistic intuitions: (1) grammaticality, (2) ambiguity (constructional homonymity), (3) synonomy (paraphrase), and (4) anomaly. However, he goes on to say:

> First, all the phenomena in question belong to a single category of 'linguistic intuitions.' Why should an entire 'science of language' be built up around them? . . . It is also difficult to imagine how the study of such 'intuitions' alone could lead to significant advances in our understanding of speech production and perception, which Chomsky himself has often implied is the 'central' consideration. (1964, p. 50 and elsewhere)
>
> Furthermore, as Householder has pointed out, it is by no means clear that making judgements about 'grammaticality'

> or 'anomaly' as such is a significant part of normal linguistic behavior. (p. 160)

To begin, an entire science of language is not built around these intuitions. Derwing's claim that it is reflects his identification of the goals of a description with the goals of a theory. A grammar must be a representation of some system of linguistic knowledge, and it should not be astounding that this system crucially includes accounts of the intuitions that characterize this knowledge. The theory, again, has as its central goal, an account of how it is possible for the human organism to acquire such a system. But in this discussion, Derwing does not even refer to this theoretical goal. Instead he introduces yet another level of requirements for a "science of language." In fact he introduces two, although he combines them: the understanding of the mechanism of perception and production, and the understanding of "normal linguistic behavior." It is not clear what is meant by the latter, although it is tempting to infer that it refers to language use. Nonetheless, linguistic theory has always kept both of these areas distinct from the goal of some understanding of linguistic knowledge. It is true that Chomsky has always established understanding of perception and production as a (not the) central goal, but he has never (including the source Derwing cites) suggested that an understanding of linguistic knowledge--reflected in part by a description that provides an account of these essential intuitions--would be coextensive with an understanding of perception and production. Rather, the claim is generally that the latter entails some aspects of the former.

Implicit in his introduction of what is really an argument about what the goals of a description should be, is Derwing's use of it as supportive of a charge of inadequacy to be levelled against the theory. But this use of the argument is inappropriate. Even if the goals of a linguistic description were defined as providing an explicit account of the abilities of language perception and production, the goal of the theory would remain: to account for the capacity of the organism to achieve these abilities. An argument about the goals of a description does not vitiate any claim about the goals of a theory, or any hypothesis about its internal structure. Derwing seeks to jump from an argument about description, to the claim that there is no linguistic theory corresponding to a (language specific) LAD. The fallacy of an argument with such a gap, as well as those we have already discussed, must not be ignored.

1.3. Response to the Criticisms

Although Derwing's arguments undermine, rather than support, his position: that linguistic theory is substantively and methodologically inadequate for the fruitful study of language acquisition, it is important to note the position, because he is not its solitary proponent. In the preceding sections, we have reviewed aspects of criticisms of the theory and proposals for alternative approaches to the study of acquisition. In the course of this review, a good number of assumptions implicit in these criticisms and proposals have surfaced. Because these are presented from diverse points of view, and are sometimes obscured by the context in which they are found, it seems help-

ful at this point to line them up in an attempt to see how they are connected. Then we may proceed to contrast them, and the position, with what characterizes a different sort of work in developmental psycholinguistics, and what defines the work proposed in this dissertation.

A productive way of treating these assumptions is to look at them from the points of view of two general questions. The first centers on what is meant by 'psychological model.' The second question involves the relationship of learning in general and language learning specifically.

1.3.1. On the Notion "Psychological Model"

When we talk about the notion of a 'psychological model,' we need to keep two in mind, and keep them distinct. On the one hand, we have a psychological model of linguistic knowledge, and on the other, a psychological model of language acquisition. Specifying the first as a model of knowledge makes it easier to avoid the error of assuming it is a model of language comprehension and production, or of language use. As a psychological model, a grammar, as it is defined in linguistic theory, is neither of these. The model of grammar must be rich enough to provide for the construction of a model of comprehension and production (linguistic processing), but it itself is not such a model, nor has ever been claimed to be one. As for its being a model of language use, a grammar is explicitly defined to exclude accounts of 'use.' That is not to say that such questions are not interesting, it is just the case that there haven't been any successful attempts to

construct a theory of language use, much less integrate it with a theory of language knowledge (cf., Chomsky 1975, Ch. 1).

We are now in a better position to describe at least what a psychological model of acquisition, defined as it is in (4), is not. It is not a model of how a child acquires the ability to process language (linguistic data), nor is it a model of how a child learns to use language as a functioning member of his linguistic community; how he learns to use language as a 'tool of communication.' However, neither is it a model of how the child breaks down the data of his environment and constructs, therefrom, a system; it is not a process model. Although Fodor, Bever, and Garrett (1974) agree that a grammar is not a model of linguistic processing, and that the disappointing results in early psycholinguistic work were due to this mistaken interpretation, and they do not interpret a model of acquisition as a model of the development of language processing, it does appear, as we saw in 1.2.3, that they have interpreted the definition of a model of acquisition in (4) as a process model itself.

Further, the criticisms that maintain that descriptions of child language must fit the data in a way that transformational grammars do not are interpreting--if implicitly--the model to be one of the development of language use. If we look back at the discussion of Bowerman, Braine, Brown, Bloom and Slobin in 1.2, this interpretation becomes clearer. Notions even so are still not entirely clear because what falls out of the overlapping notions under the label 'use' are two distinct ones.[6] The first is <u>pragmatics</u>[7] and the second, <u>semantics</u>. Although the following may appear to be overly simplistic as a

characterization of the two, it is not, I believe, inaccurate. The pragmatic view holds that a child's hypotheses are based on what he needs to say, and learns that he needs to say, to become a communicating member of his linguistic community. In the second view, the structure of the child's linguistic hypotheses is controlled by what he means to say.[8] It is not hard to see how these two notions fell together when the goal of a descriptive model in language acquisition became providing an account of what the child says and means, and as we saw, this did become the goal.

A question that needs to be asked is what is involved in the development of the system of linguistic knowledge. In order to ask such a question, we must be able to do two things. First we must be able to abstract away from descriptions of child speech. The grammars we infer for children in the course of their development will thus not be reflections of what they say, but will be substantive guesses about their linguistic competence at any one point. Second, we must be able to formulate questions about a developing linguistic system that are distinct from questions about the development of other systems of knowledge. Ideas about the separability of such questions will necessarily affect the picture we have of a model of acquisition. Related to these ideas are the assumptions we have placed in the second category--those that underlie views about the relationship of learning in general to language acquisition specifically.

1.3.2. On the Question of "Learning"

More than a few investigators contend that the mechanism of language acquisition is isomorphic with some general learning mechanism,

and that there is nothing that we might call a "linguistic" learning mechanism. This notion runs parallel to the view of language that holds there is no linguistic system distinct from other systems of knowledge. It seems plausible that there are many factors, and not a few of them non-linguistic, that determine what a child may understand and talk about. What seems implausible is to deny that there are linguistic factors which are crucial in determining how the systems that describe the language children understand and produce are constructed. We have seen, nevertheless, that investigators such as Derwing maintain a position that rests crucially on just such a denial. For Derwing, there is no justification for any system that seeks, as a central goal, to account for the shared intuitions of a community of adult speakers of a language, nor by extension, is there any inherent language specific learning mechanism that mediates the development of the system.

Suppose we were to maintain this position from the beginning in an attempt to determine the internal structure of the language acquisition device. It becomes quite difficult to propose answers to questions about how it is that children pay attention to precisely certain kinds of information and develop precisely certain kinds of systems. Even if we ignore the claims about the relatively inferior quality of the data that the child gets and the relatively short time it takes for a child to develop (nearly) adult competence, we can't ignore the observation that speakers in a linguistic community each have developed some system that does allow them a set of linguistic intuitions which they all share, by and large. While there is suffi-

cient evidence that such intuitions are real: speakers can discriminate between grammatical and ungrammatical strings, they are generally sensitive to ambiguity (constructional homonymity), synonomy and anomaly (cf. p. 37 above), there is little evidence to suggest that any of these is ever made explicit for children by parents. In fact, the principles accounting for the intuitions are of a sufficiently complex nature to make it unlikely that adult speakers are ever conscious of them.

We needn't restrict ourselves to questions about the possible source of the system that controls these intuitions to see that denying the possible existence of some language specific learning mechanism at the outset makes it difficult to provide substantive answers. Even a fairly gross level of observation allows us to see this. Neurologists have long noted that for most normal right-handed individuals, the "language center" of the brain is largely localized in the left hemisphere. Such an observation suggests that language belongs at least to a subset of possibly related cognitive systems--those also localized here--as opposed to belonging to some as yet unspecified general set if we are to be able to draw any conclusions other than that this systematic localization is accidental.

Further, there are findings from two sources that enhance our claim that there is a language specific mechanism separate from a more general cognitive system. In one discussion of her study of Genie, Curtiss (1979) notes that certain aspects of Genie's cognitive development during the time she was observed exceeded her linguistic development, which remained "frozen" at what Curtiss describes as a fairly

immature state. The three areas Curtiss discusses in this regard are tool use, drawing and the construction of "mental maps." In all of these, Curtiss suggests that Genie exhibits fairly sophisticated behavior, while her observable language lags.

The second source is a discussion of Greenfield and Schneider (1978) by Curtiss, Yamada and Fromkin (1979). In the first study, it is claimed that certain cognitive-motor functions are prerequisite to what are seen as parallel functions in the linguistic domain. In particular, nesting activity and hierarchical organizing activity are claimed to underlie embedding and hierarchical structure in grammars. Curtiss et al. show, however, in experimentation with retarded and normal individuals, that activity in either domain is independent of the other. In particular, nesting activity, expressed by a task involving seriation of cups, had no relation to linguistic nesting--i.e., embedding. Subjects who were highly successful at the task showed little or no embedding, and other subjects whose performance at the task was poor exhibited quite sophisticated instances of embedding, involving the interpretation of "gaps." Equivalent results were found for the relationship of the task involving hierarchical organization to linguistic hierarchical structure.

None of these findings, either in the case of Genie or in the case of these experiments in the "activity" domain would have a plausible explanation were we to deny the existence of a language specific mechanism. We must assume a mechanism specific to the acquisition of language, then, distinct from a more general mechanism controlling development in all cognitive domains.

Beside the position characterized by the denial of the existence of a language specific LAD, there is the view we have seen that emerges in the work of Slobin, Bowerman and Bloom. This is that the primitives of a model of language acquisition are the cognitive constructs which control the development of semantic and pragmatic knowledge. As we pointed out (cf. p. 21, above), the problem with making this assumption at the outset is that if it is acknowledged that a transformational generative grammar (in some form) is the end result of acquisition, we are left with questions difficult to answer about how transition is made from the proposed primitives to the sorts of principles that govern the construction of grammars--the principles of Universal Grammar.

The argument here is not that such questions may not turn out to be appropriate ones to ask, in the event we do discover that such a transition exists. Rather, the argument is that making such assumptions about the shape of the language acquisition device prior to determining information about its substance or without reference to the developed system (the grammar) which is presumably its output will make it difficult to answer some questions and may remove the possibility of posing others that may be important. No one would argue that there are no interesting questions about the different processes involved in the child's analysis of the linguistic data in his environment, or in his analysis of the relationship of those data to the environment. But to date, there has been little progress toward establishing any set of principles for how adults with full linguistic competence may deal with actual data--how they process language, or

how they relate language to their world. To start here in an investigation of child language would render useless most of the knowledge we have acquired about what must be included in a description of adult linguistic competence.

I merely wish to make the assumption that we can look separately at that part of the LAD that crucially involves the principles of Universal Grammar. Of course, this involves the presupposition that these principles are part of an inherent LAD. Chomsky (1975, cf. especially Ch. 1) argues convincingly for the plausibility of the assumption and its underlying presupposition. But even in the absence of the type of argument Chomsky proposes we can look to questions of learnability to justify the making of the assumption that there is an inherent LAD. We have accepted the cited observations that children do not receive information about ungrammatical sentences of any crucial kind, but may in fact hear various ungrammatical strings interspersed with grammatical ones (Brown and Hanlon 1970; Snow and Ferguson 1977). Gold (1967) demonstrated that given the absence of such negative data, if we assume that the child's learning mechanism has only some general induction procedure, learning of the language (in the sense of construction of the "right" grammar) will be impossible in the child's lifetime. Thus we would have no hope of accounting for the child's ability to construct most of the grammar by the time he is 3 or 4 years old, if we did not assume that there is some fairly rich innate specification of the set of plausibly right guesses for a grammar to limit the range of possible grammars. That this specification is of the form of an LAD, the requirements of which were outlined in (4)

requires an additional assumption, but one that is now doubly supported. Assuming such a framework has allowed the most progress toward the determination of adult grammars. No progress toward any unified descriptions of adult linguistic competence has been made in the absence of this assumption. Moreover, the more specific to linguistic phenomena we assume the LAD to be--that is the closer we take it to be to the principles of Universal Grammar--the more limited is the range of "guesses at grammars" allowed the child, and the closer we come to a realistic picture of how it is possible for him to construct a grammar in the relatively short time that he does.

Nevertheless, we find such linguists as Derwing contending that maintaining that the related position that a child comes to "the language learning situation" fully equipped with some specific set of "innate linguistic universals" which makes possible the acquisition of a "rich and highly specific system" defines away the problem of determining the substance of the acquisition mechanism (p. 70). In reality, the position outlines the problem quite clearly, and provides precisely the narrow field that we require to set questions for which we have a better chance at answers. In the next section we will look at the sets of questions that can be posed in such a narrow framework and make precise the paradigm question that forms the basis of this dissertation.

Before examining these, we might look at some possible problems that face a research program which assumes it can look separately at the development of a grammar within the context of some innate mechanism. The first problem deals with the claim someone might make that

the grammars we infer as reflective of a child's linguistic competence at any point are really artifacts of some other phenomenon operating. As we have stated, there are several factors affecting the linguistic competence. Hence the question arises of whether (the consequences of) such factors can be filtered out of the data that are considered to be reflective of the child's linguistic competence. In fact, making the assumption at the outset that we can consider the child's grammar separately enables us to ask this very question. It would not come up if we were not concerned with the possibility of identifying and studying an LAD distinct from acquisition devices for other cognitive functions. Further, given the assumption that we can pursue such a study, it is not so difficult to establish how we might proceed in determining whether we are correct in basing inferences about grammatical competence on particular data.

As examples we can take the "two-word stage," and the construct MLU (mean length of utterance, measured in morphemes, cf. Brown [1973]). MLU is proposed as a metric by which children's linguistic sophistication can be classified. There is quite wide agreement now that the limit to two words is not reflective of linguistic competence: Children at the "two-word stage" understand utterances that are longer than two words (Bloom 1973), and thus the limit on speech length is controlled by some other mechanism. We would not, then, want to infer grammars whose shape is determined by the length of such utterances. Then how should we view a construct such as MLU? It may be considered to show a better correlation with children's speech patterns at various stages than does age, as Brown has proposed. But, given the observa-

tion that children's comprehension generally exceeds production, we would not want MLU to determine the shape of grammars that reflect children's linguistic competence. The question is, then, if we are going to consider the development of the child's grammatical system given some innate mechanism, what will control our inferences about the shape of these grammars? They will depend largely on our guesses about the requirements of the innate mechanism, which will mediate our use of speech data.

If we take (4) as descriptive, in a very general way, of these requirements, and if we add the substantive descriptions of natural language within the framework set by (4), the grammars we infer for children could never be determined by length of utterance because nothing in our findings for natural language leads us to conclude that some fixed length of utterance has any bearing on structural descriptions.

A subtler interaction of the data and the theory in a second example gives us another perspective on the problem of the separability of questions dealing with linguistic theory in the course of acquisition. Slobin (1966) observed that children who were unable to understand a sentence such as (5a) were able to interpret strings such as (5b) correctly.

(5) a. The dog was chased by the cat.
b. The sandwich was eaten by the policeman.

It is generally suggested that the child's performance here is due to his ability to establish the relationship between <u>sandwich</u> and <u>policeman</u>: policeman eat sandwiches, but not vice versa. This is the case

even though the child at the relevant point in linguistic development does not control the grammatical construction passive (cf. Roeper 1978). We would want, thus, to be sensitive to the child's use of these inference procedures, and not base a guess about a child's linguistic competence solely on his performance with a string such as 5(b). However, care must also be taken to recognize that such procedures operate in a limited way--Roeper suggests that their "use" is restricted to unknown structures. Hence, a child at the stage reflected in (5), who has achieved control of basic clause word order would not use such an inference procedure to interpret (6) incorrectly. He would understand it as nonsensical.

(6) The sandwich ate the policeman.

Such findings suggest further that a model of the development of grammar is not isomorphic with a model of the development of some model representing pragmatic competence. And again, we cannot get any productive questions out of a research program that assumes as primes the cognitive constructs that underlie conceptualization of pragmatic facts such as those reflected in 5(b). In such a framework, there is no explanation of how, given a general learning device, such constructs would lead precisely to the grammars that have been shown to be required for a descriptively adequate account of adult language competence. We have no explanation for the mapping from these constructs onto the developing syntax.

It appears then, that these examples do not in fact pose problems for a research program which assumes that the development of grammar is at least a logically separable problem within the context of some

innate hypothesis mechanism. Rather, the assumption allows us to consider phenomena such as mean length of utterance, or interpretation of sentences such as (5) with a better view of their relationship to the developing grammars we infer (that is as long as we avoid allowing our goal to become accounting for [every aspect of] child utterances). The grammars that we guess at will be subject to a theory of grammar. Hence we will have some principled basis on which to decide whether an observed phenomenon reflects grammatical competence at some point or is the product of some other aspect of cognitive (or even, perhaps, motor) development. Of course, there may at times be indeterminancy about the source of some phenomenon we observe in child language, but if we assume that the grammars we infer for children are constrained by whatever principles make up the LAD, we will have a way of evaluating our guesses, and the indeterminancy will be somewhat reduced.

1.4. Current Approaches to the Acquisition Problem

We have seen, in the foregoing sections, that for the most part, the failure of developmental psycholinguistics to make much progress toward an account of language acquisition is traceable to its move away from a notion of LAD of the sort defined in (1) and outlined in (4). The sources of this move, we have seen, have been either the assertion that although a grammar may be the end point of acquisition, its development is essentially a function of the development of other mental abilities, or that a generative (transformational) grammar is not the end-point of acquisition, and thus there is no need to posit an inherent mechanism such as an LAD to account for language acquisi-

tion. In both cases we are left to observe child language with little hope of proposing any account of the relationship of developing language to (full) adult competence. This final section focuses on certain of the most promising types of research programs that can be carried out given the assumption of an LAD outlined in (4), included within which the work in this dissertation is carried out.

The strongest claim about the substance of the LAD we can make is that it is (essentially) identical with the principles of Universal Grammar (UG). Chomsky (1980) makes this claim explicit as he has in several other places (1965, 1973, 1975, etc.). A good deal of progress toward an understanding of adult competence has been made by concentrating on the specification of UG. This is the essence of the research program outlined by Chomsky. We have already referred to the fruit of work within the bounds of syntax, in this program (cf. 1.2.4): the proposal of principles which limit the shape of base rules (Jackendoff 1977; Selkirk 1977; Hornstein 1977), those limiting the transformational component to the simple rule, "move constituent," and those establishing conditions on the interpretation of referring expressions such as pronouns, anaphors and quantifiers. These principles define a (fairly small) finite set of grammars from which the 'correct' grammar can be selected (constructed) on the basis of particular data. The selection procedure so defined assumes the presence of "parameters." For example, a principle of UG may state that the form of a phrase structure rule must be that in (7).

(7) $C^n \rightarrow \{\text{Specifier}, C^{n-1}\}$

(7) states that category expansion requires a specifier and the category head. The parameter along which languages may vary is the order of the specifier and head. Only the appropriate data (Chomsky's "experience") will determine in which order the specific grammar will generate the constituents of the category. Hence, for a child, the environment furnishes the data specifying the correct ordering of the constituents while (7) is claimed to be included in the LAD.

Within the framework of the claim that principles of UG constitute the LAD, the questions we can actually ask about language acquisition come into focus. It is significant that the greater part of our knowledge about the LAD has come about as the result of research focusing on the description of phenomena in "adult" language. For the most part, as we have seen, it has not been possible to make much progress by looking at child language because the studies that did look at it were not carried out within a framework which made the logical distinction between the effects of the LAD and the contributions of other factors to the development of linguistic knowledge. This claim about the LAD, however, together with the claim we have made that it is possible to study the development of syntax separately from the development of other mental capacities, allows us to turn back to child language with some interesting questions and with the possibility of productive answers.

1.4.1. Longitudinal and Experimental Studies: Evaluating Assumptions and Claims

There are a growing number of child language studies which accept, as basic assumptions, that the end point of language acquisition is

correctly described by a generative grammar, and that points in the course of language development can be characterized as immature forms of these grammars (Fischer 1971; Sheldon 1974, e.g.). Moreover, there is a growing number of studies that explicitly assume as well the presence of universals and question their role in acquisition. These fall into three major categories. The first can be characterized by work reported in Lust (1978), Erreich et al. (1978), Hamburger (1978, 1980), Roeper (1972, 1977, 1978), and Goodluck and Solan (1978).

In these studies there are attempts to determine experimentally whether specific principles of UG are present in the developing grammars of child language (Lust, Erreich et al.). Hamburger (1978, 1980) has a slightly distinct approach. Here, an attempt is made to determine the limits placed on child grammars by the principles of learnability (Hamburger and Wexler 1973; Wexler, Culicover and Hamburger 1975). Although these are determined outside of a descriptive framework, within a learning theoretic framework they have been found to relate to the principles of UG in important ways (Wexler, Culicover, 1980; Wilkins 1980).

We will focus, for the moment, on the work by Roeper and the work in Goodluck and Solan as representative of the approach in this first category. The work is important because it reveals a principled attempt to limit the kinds of grammars we induce from actual child language data on the basis of principles that have been found to be suggestive in the attempt to constrain the kinds of grammars that are proposed to account for aspects of adult language. There are problems of detail to be resolved in connection with this work, but it is

important to note that they are not systematic weaknesses in the framework adopted. The problems show, rather, how difficult it is to deal with data from child language, even given a fairly constrained theoretical framework and an apparently fruitful experimental paradigm.

We have already raised the question of how to determine what features of child language we should try to account for (directly) in grammars, and we have seen that the theory, along with what we have discovered about natural language given the theoretical framework, leads us correctly to ignore some--apparently systematic--aspects of child language in our induction of grammars. However, there are some more subtle aspects of child language that are not as easy to make decisions about, and it is such aspects that cause problems. An analogy from the development of another faculty that is generally accepted to be unlearned may make the dilemma somewhat clearer. Like linguistic competence, the ability to walk develops slowly and in the course of its development there are many instances of imperfect walking-like behavior. Do we interpret this behavior as the result of an immature innate system or as the result of other mechanisms, not specific to walking: developing muscle tone, incomplete neurological conditions and the like, interfering with our ability to see the innate system? While there may be no easy answer to this problem, there do seem to be ways to get closer to the innate system underlying language acquisition--to recognize when we <u>are</u> catching a glimpse of it, or when it is obscured from view.

Matthei (1978) conducted an experiment in which he could check children's intuitions about the interpretation of the reciprocal

anaphor each other. He was looking specifically for evidence of the specified subject condition, a principle of UG constraining the operation of rules of interpretation (Chomsky 1973, 1980, here restated as the Opacity Condition), in children's hypotheses for a rule of each other interpretation. The Specified Subject Condition (SSC), which we can see operating in (8) blocks the interpretation that for each parent, the child(ren) want the parent to like the other child.

(8) The children want their parents to like each other.

Matthei tested children's determination of the references cf each other in embedded clauses such as those in (9).

(9) a. The chickens said (that) the pigs were hitting each other.
b. The cows want the lambs to kiss each other.
c. The pigs noticed the boys patting each other.

In 9(b) and (c), the Specified Subject Condition precludes the rule of interpretation from establishing coreference between each other and cows and pigs respectively. In 9(a), not only does the Specified Subject Condition operate, but another principle of UG (which we will consider in more detail in the following chapters), proposed in Chomsky (1973) and subsequently redefined, the tensed-S condition, operates as well to block this interpretation.

In Matthei's results, 64.4% of the responses involved the children interpreting sentences such as (9) such that the matrix NP and the embedded NP were a conjoined antecedent of each other: the chickens and pigs were hitting each other, the cows and lambs kissing each other, and the pigs and boys patting each other. Complement type had no apparent effect on these results. Matthei interpreted the

results as indicating that the Specified Subject Constraint was not present in the LAD and indeed that such constraints are learned (pp. 164-65).

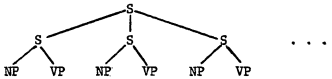
The immediate question is not whether the constraints are or are not present. Rather the question is whether these results lead us to conclude that they are not. The answer here must be, I think, no. The children's interpretation does not constitute a direct violation of the SSC. This fact is apparent when we compare their interpretation of (9) with the interpretation we discussed of a sentence such as (8), which is explicitly blocked by the SSC. There is a good possibility that the "immature linguistic behavior" (following our analogy of "immature walking" behavior) we are observing does not directly reflect the operation of the LAD. Rather, its operation is being interfered with by some other factors. Matthei himself observed that in simple sentences children more readily interpreted conjoined NPs as antecedents of each other than they did simple plural NPs.[9]

This preference for antecedent type undoubtedly has a major effect on children's interpretations of the anaphor in complex sentences as well. While it is not possible for us to conclude that this hypothesis about the antecedent requirement of each other blocks the operation of the Specified Subject Condition, it does seem likely that it provides enough interference to make it equally impossible to conclude that the SSC is completely absent: our view of the LAD is obscured by the presence of other, more peripheral hypotheses. We could speculate about the source of this particular hypothesis. It may be that for perceptual reasons, conjunction provides a more

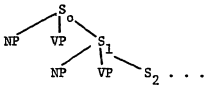
"secure" signal about plurality than does a single morphophonological cue; hence a preference for a conjoined antecedent for an anaphor that requires a plural. On the other hand, for semantic reasons, having to do with the reciprocal interpretation of <u>each other</u>, a conjoined NP may be a more secure source of separable plurality than a simple plural NP.[10] Testing with anaphors that require antecedents of either type might serve to discriminate between these two proposals. At this point both of these proposals are speculative. Nonetheless, they seem plausible enough (and certainly testable) to suggest accounts other than the absence of SSC that take steps toward an explanation for children's interpretation of (9).

A second problem that exists in these studies is that of determining what is appropriate for us to infer as a hypothesis in the grammars of children. Tavakolian (1977, 1979) and Solan and Roeper (1978) propose two hypotheses for children's analyses of embedded sentences; the conjoined-clause and S-initial hypothesis respectively. We see the structures assumed in these in (10) and (11).

(10) Conjoined-Clause (Tavakolian 1977, 1978)



(11) S-initial hypothesis (Solan and Roeper 1978)



Tavakolian proposed this structure to account for children's interpretation of the missing NP in sentences such as those in (12).

(12) a. The sheep that knocks down the rabbit_____ stands on the lion.
b. The sheep knocks down the rabbit that_____ stands on the lion.
c. The sheep knocks down the rabbit_____to stand on the lion.
d. The sheep knocks down the rabbit and_____stands on the lion.
e. The sheep tells the rabbit_____to stand on the lion.

The results of Tavakolian's experiments indicated that children exclusively interpreted the subject of the matrix clause as the controller of the missing NP. In certain cases this interpretation will coincide with the facts that are reflective of an adult grammar, while in other cases (such as (b) and (e), for example) it diverges. As (d) indicates, the "first" (=left-most) NP--the subject of the matrix clause--is indeed the controller in conjoined structures. Hence, Tavakolian proposed (10), and the schema in (13) that indicates the structure off of which interpretation is done.

(13) $[_S\ [_S\ NP\ldots\ V\ldots\ NP\ ]\quad [_S\ \Delta\ldots\ (to)\ldots\ VP]\ ]$

Roeper's S-initial proposal has the same consequences: it accounts for the same range of interpretations as does Tavakolian's conjoined-clause hypothesis. In this respect, it could be thought of as no more than a notational variant of Tavakolian's structure. The difference lies in the fact that the conjoined-clause structure is posited precisely to account for children's interpretations of missing NPs in complex constructions, while Roeper sees this interpretation as

a particular case of a more general phenomenon. He claims that children attach all new material to the highest S.[11] Roeper maintains that this hypothesis then provides an account as well for the fact that children at first place negatives and quantifiers in sentence initial position (and still interpret quantifiers not in S-initial position as having the entire S as scope), for the fact that their grammars generate questions such as Did you can come?, and, finally, for the fact that some children sometimes interpret the complements to perception verbs (as well as participial complements to other verbs) as subject controlled (Solan and Roeper 1978; Goodluck and Roeper 1978). These phenomena, however, have alternative accounts which fit into a framework without the S-initial hypothesis and which we will consider in more detail in the following chapters. The justification for the S-initial hypothesis then, as opposed to the conjoined-S hypothesis, for example, cannot rest on the account it provides for these phenomena.

What is striking about both of these hypotheses is that neither is connected crucially to any claims that come from the principles of Universal Grammar about how a grammar should look. Very simply, there is no universal that suggests that an optimal grammar is one in which all clausal recursion is reduced to clausal concatenation. This absence should make us cautious about proposing either of these as a hypothesis about the shape of the developing grammar. With respect to embedding, Roeper correctly criticizes various proposals for early acquisition systems that are based on the premise that a simple system (such as one that depends exclusively on grammatical relations, for

example) simplifies the acquisition task. He points out that (if the assumption is that later systems do depend, in a systematic way, on earlier ones) such proposals make the task more complex, in fact, because (as we have pointed out) they leave unanswered questions about what might be the connection between these quite simple systems and the complex systems that are required to account for higher level competence.

On the other hand, Tavakolian proposes that a feature of the conjoined-clause hypothesis is that it simplifies the acquisition problem. She reasons that the hypothesis is a universal feature of acquisition; "that all children will initially analyze multiple-clause sentences as conjoined clauses" (p. 82). Thus, the number and kind of hypotheses that a child can select from in the course of acquisition is greatly limited; a positive result. It is true that simplification of the child's task by limiting the hypotheses available to him is a positive result. But this is true only in a narrow sense. The principles that limit the child's hypotheses simplify his task if they bring him closer to the grammar that reflects adult competence--the endpoint of acquisition. In general such principles are those that are included as part of UG. A principle proposing that all children's grammars initially generate all multiple clause structures by iteration of the initial (root) S (or, incorporating Roeper's alternative, that all new material--including clauses--is first attached to the initial S) is not related to any principle of UG, any more than is the conjoined-clause hypothesis. Although a grammar corresponding to this description might appear simpler in some way, proposing it does nothing to simplify

the acquisition task. In fact, it may introduce complexity since having proposed such a stage, one must outline the framework that would provide the child with the data necessary to reformulate structures such as these.

Suppose we look specifically at the acquisition task as well as at the theory. The question about simplification is significant here, too. One of the tasks that face the language learner is to establish the strict subcategorization facts about lexical categories in his language. The question about whether some lexical item allows a clausal complement entails the knowledge that the syntactic category of which that item is the head allows clausal recursion. If children seem to exhibit knowledge about subcategorization at some stage, it is plausible to infer that their hypothesis mechanism is able to determine the range of syntactic categories that can occur as complements to any other syntactic category. The range should include whatever category defines the clause.[12]

Roeper provides evidence that at the stage during which he proposes the S-initial hypothesis is operative, children are conscious of the subcategorization requirements of the lexical item 'put.'[13] Hence, we can infer that their grammars generate as possible complements NP (N^{max}) and PP (P^{max}). If S is a member of the set of syntactic categories, it seems exceptional to assume that the grammar will not include the hypothesis that S is a possible complement as well and that lexical categories may subcategorize for it. In fact, the simplest base configuration for a learner to filter data through seems to be something like the following, which is close to what is being found to be required for UG.

(14) a. $C^n \rightarrow \{(\text{specifier}) - C^{n-1} - (\text{compl})\}$

b. $\text{Compl} \rightarrow C^n$, where $C^n = \begin{Bmatrix} N^{max} \\ V^{max} \\ P^{max} \\ A^{max} \\ M^{max} \end{Bmatrix}$

A plausible scenario of the learning would involve establishing the range of complements for any given lexical category. Given a schema such as (14), and the available data, the child could arrive at the complementation structure for any lexical category. Both the conjoined-clause and S-initial hypotheses require the child to go through such a procedure but, in addition, require him/her to unlearn the restriction on embedding that would be part of their grammar. Thus the hypotheses would not simplify the learning task in any way.

Of course, rejecting the S-initial and conjoined-S hypotheses leaves unexplained children's interpretations of (12). However, recent work by Hamburger and Crain (1979) suggests that the experimental method employed by both Tavakolian and Roeper leads, in some cases. to the interpretation of the relative clauses tested as non-restrictive, rather than restrictive. Analyses have been proposed (Emonds 1979) suggesting that the correct account of non-restrictive relatives involves sentence conjunction. Hence, in a number of cases, the child's interpretation of the relative clause as a conjoined clause would be the result of his interpreting it as a non-restrictive, rather than the result of the conjoined clause hypothesis limiting the shape of his grammar.

What we must conclude from the problems we have discussed in this work and Matthei's is, again, not that the framework for looking at child language is incorrect, but that, rather, it requires a tight interpretation. In the case of Matthei's results, we noted that given the precise definition of the Specified Subject Condition, children's responses could not be interpreted as showing that this constraint is absent from the child's hypothesis mechanism. The case of the conjoined-clause hypothesis and S-initial hypothesis suggests that if we look at the theory (as well as at the requirements of acquisition), the principles we propose as constraints on the shape of children's grammars will look more like the constraints that have been proposed as part of UG, and that will permit children's grammars at various points to be less far from the endpoint of acquisition.

1.4.2. Formal Models of Language Acquisition

The second category into which work asking questions about the precise role of universals in language acquisition falls, we will exemplify with early work of Gold (1967) and the work of Hamburger and Wexler (1973), Wexler, Culicover and Hamburger (1975, 1977), and Wexler and Culicover (1980), although there are a number of other studies along these lines. For a review of such work, see Pinker (1979). The work is removed both from the area of child language data, and from the area of descriptive linguistics, though it has important implications for both, in that it deals directly with the question of the learnability of linguistic systems.

Gold (1967) first established that an important question to ask

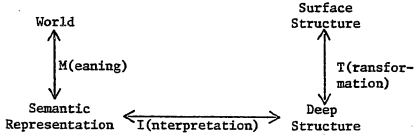
was whether a class of languages was learnable, not just a given language. Second, his proofs led to the conclusion that no set of languages (context-sensitive, context-free, finite state, and primitive recursive languages) are learnable if only positive data are available (what Gold called a text situation). Negative information about what strings are not in the language is crucial (Gold's informant situation). Another result of Gold's work, however, is that apparently the general class of languages computable by a Turing machine is not learnable in even an informant situation, hence certain transformational grammars not constrained in the appropriate ways may not be learnable at all (cf. discussion in Matthei 1978).

Gold's work, however, was done without the assumption that grammars must be constrained, even once the class of languages is defined. In the work of Hamburger et al., attention is paid precisely to the formal requirements of learnability. That is, constraints on the form of grammars are derived directly from what is required to prove learnability of the class of transformational languages, and in fact to prove learnability given only a 'text' situation. In this work, however, this situation is more explicitly defined, and the particular definition forms one of the assumptions basic to the general program of research. The learner has access to the base structure and the surface structure of any sentence he is presented with. These structures are, in the learnability proof, considered as ordered pairs (b,s). The general learning procedure is for the learner to hypothesize some transformation, T, that relates a given pair (b,s), and then to test T on some new datum to determine if the grammar with it

will generate the new datum. The goal of the learnability proof is to establish the learnability of the class of transformational grammars (and, thus, any <u>one</u> of them) given this general learning procedure. The proof requires the postulation of a number of constraints on the notion "transformational rule," and these constraints, such as the Binary Principle, the Raising Principle, the Freezing Principle, the No Bottom Context Principle, etc., are the best known results of the learnability work. Before explaining these results, and their relationship to parallel results--the discovery of constraints on the notion "transformational rule"--in the descriptive domain we will look more closely at two of the assumptions central to the learnability work.

The first assumption underlies the claim that the learner has access to the (b,s) pairs. The way that the learner has access to the base structure (b) is through the interpretation of the (surface) sentence. Hence, the assumption is that a base structure for a given sentence (sentence here is used with rough equivalence to <u>surface structure</u>) is uniquely derivable from the meaning of the sentence. Culicover and Wexler (1974) schematize the structure of this assumption, and the closely related ones it entails, as (15)

(15)



("The Semantic Basis of Language Acquisition: The Invariance Prin-

ciple as a Replacement for the UBH" SSWP #50 UCI 1974, p. 5)

"World" and "Surface Structure" are taken to be available for direct observation, while "Semantic Representation" and "Deep Structure" must be constructed. The acquisition task, therefore, is seen to involve the learning of the mapping procedures M, I and T, and, the emphasis of the proof is to establish the learnability of the mapping procedure T. The notions of M and I are crucial to the view of the mapping procedure T, and these are somewhat problematic.

Culicover and Wexler (1974) acknowledge certain problems. They point out first that the procedure for M may be incompletely developed; that the learner may draw the wrong inferences about the Semantic Representation, and thus that M deserves some attention. The greater attention, however, is paid to the relation I, which supposes that a universal set of Semantic Representations is mapped onto some universal base. It is allowed that a Universal Base Hypothesis is too strong, since it removes the possibility of variation in constituent order across languages, and Culicover and Wexler propose, in its place, the Invariance Principle that places constraints on how constituents may be ordered by base rules.[14] The Invariance Principle is claimed to be sufficiently weak to account for variation across languages, but sufficiently strong to restrict the class of possible realizations of I so that it is readily learnable, and available to the learner for the task of learning T.

Still remaining, however, is the assumption that the base structure is determinable from the semantic representation. And this

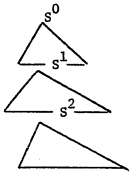
assumption is what is most problematic because it is not an assumption shared by the research program in the descriptive framework. The assumption has been rejected since Chomsky (1970) where it was first argued that certain aspects of semantic interpretation (focus, presupposition, scope of quantifiers) had to be determined from surface structure, to the present, where it is widely accepted that (an enriched) surface structure is the exclusive input to rules of semantic interpretation. In fact, the arrow labeled I in (15) could be rotated 45° counter-clockwise to indicate the appropriate relation between semantic representation and surface structure in the descriptive framework.

A second assumption, which was problematic in the early Learnability work was that the learner has an infinitely large set of grammars from which to construct the grammar for his/her language. This assumption was valid given a framework such as (4) for which it would be reasonable to motivate constraints on the notion 'transformation' by way of a learnability requirement. The assumption became problematic, however, when the goals of linguistic theory shifted such that the descriptive requirements there pointed to constraints limiting the class of grammars to a (fairly small, in principle) finite set of grammars. This reduction in the class of possible grammars is claimed to reduce the learner's task, and therefore to remove the need for a proof of learnability (cf. Chomsky 1980).

There are two changes that characterize the evolution of the Learnability Theory and which respond to this problem. Both of these changes derive from the change in the Learnability requirement. The

earlier work (Hamburger and Wexler 1975) showed that transformational grammars were learnable given a limit on the complexity of data required to reveal an error. This limit was expressed as the Boundedness of Minimal Degree of Error (BDE). Degree 2 Learnability is proved. In Wexler and Culicover (1980) this limit is lowered to 2, and that is, the Learning Procedure can hypothesize the right set of transformations from data no more complex than (16) (where depth of embedding is a measure of complexity).

(16)



First, constraints on grammars required for error detectability on Degree-2 phrasemarkers (thus allowing Degree-2 Learnability) themselves turn out to result in a finite set of grammars (for discussion cf. Wexler 1980). Second, and of some significance is the centering of attention not just on the learnability of a grammar (or the set of transformations), but the <u>feasibility</u> of the theory of the grammar. Feasibility relates to questions of human capacity, and directly involves responding to problems of the complexity of data required for the learning of a grammar. Thus, if there are limits on the complexity of data that a child can handle, the simpler the data required for a grammar to be learnable then the more feasible is that grammar.

Wexler and Culicover (1980) attribute the feasibility requirement to Chomsky (1965), and go on to state that

> In our terms, feasibility may be called 'easy learnability,' that is, learnability from fairly restricted primary data, in a sufficiently quick time, with limited use of memory. (p. 19)

They note that feasibility is coextensive with explanatory adequacy (where a theory is explanatorily adequate if it selects a descriptively adequate grammar on the basis of primary data [Chomsky 1965]) if in the definition of explanatory adequacy, primary data is stipulated to include precisely what the child has access to. The reason for the importance of a criterion such as feasibility is that even if the class of possible grammars were reduced to two in some theory, if the choice between them were to depend on data too complex for (or for other reasons not available to) the child, then the theory generating the grammars would be inadequate. Thus, it is maintained that the finiteness of the set of grammars is not a sufficient result.

The first assumption, that input to the learner includes (b,s) pairs and that the base representation (b) is determinable from the semantic interpretation, is still somewhat problematic, although it is addressed in Wexler and Culicover (1980), with specific reference to trace theory (Chomsky 1973, 1975, 1977; Fiengo 1976). They suggest that the learner may be able to infer (b) on the basis of correct inferences about the position of the trace (t_i) in the surface string, its relationship to some NP_i elsewhere in the string and the semantic interpretation of the string. The admitted problem is to formalize how this inference might proceed.

Dealing with the relationship of the assumptions made in the framework of Learnability Theory to results that have led to the evolution of Linguistic Theory in the descriptive framework is important since a goal for learnability is also to define the class of descriptively adequate grammars:

> In the work to follow we will demonstrate once again that considerations of explanatory adequacy provide 'one of the main tools for arriving at a descriptively adequate grammar.' We will do so by showing that the criterion of learnability leads us to hypothesize that natural language has certain properties, properties we can then test descriptively. To the extent that these tests support the hypothesized properties, the criterion of learnability will have been a tool in attaining descriptively adequate grammars. (Wexler and Culicover 1980, pp. 21-22)

In fact, one of the very positive results of the learnability work, the Binary Principle, is considered so because of its striking similarity to Subjacency, a constraint on the functioning of movement transformations first proposed in Chomsky (1973). Roughly stated, the Binary Principle insures that the structural description of any transformation may apply only to a node in a phrasemarker in the cycle of the transformation or in exactly one cycle below it. Such a constraint is required to prove learnability of the class of transformational grammars by the learning procedure assumed in the Learnability work, because as we have seen, it must be shown that an error in the formulation of a transformation made by the learner is detectable on some phrasemarker that is not arbitrarily large, that is, BDE must hold. The Binary Principle is required as well in the Degree-2 Theory. Subjacency restricts the operation of major movement rules to a single cycle or to adjacent cycles. Thus, we see that the descriptive con-

sequences of the two constraints are equivalent.

There is a problem, nonetheless, with the relationship of the learning theoretic basis for the Binary Principle and the descriptive (syntactic) data that contributed crucially to its final formulation. The problem is that apparently the final formulation of the Binary Principle, while syntactically adequate, is not adequate for the requirements of the learnability theorem proof. Culicover and Wexler (1977) note this, however, and propose steps that could be taken to rectify it (cf. their fn. 54). There are problems too, from the descriptive point of view, with certain of the other principles, in particular with the Freezing Principle, and possibly with the Raising Principle. Required for the proof of the theorem, these principles are being found to make incorrect descriptive predictions (for some discussion, see Williams 1980).

Even in light of the problems we have outlined,[15] the assumptions about grammatical theory on which the learning procedure, and hence the proof seem to be based and the relationship between the learning theoretic and descriptive domains of the constraints derived in the proof, the learnability work is important, because it reveals new ways of looking at the connection between the notions about acquisition and Linguistic theory. Suppose, for example, that the constraints derived in the learnability proof were to serve as a sort of evaluation metric along which a theory of markedness could be worked out. The most highly valued grammars would then be defined as those that obeyed the constraints, and the learner would hypothesize these first. Thus, the number of available grammars in the absolute would not be reduced, but

their distribution would be skewed in favor of learnability. Lasnik (1980) points out that such a result in the effort to constrain a theory of grammar is a positive one from the point of view of acquisition.

Even the problems we have discussed raise themselves possible lines of inquiry. It does not seem impossible to maintain a learning procedure for the transformational component that allows the learner access to base and surface structures without assuming that the access to the base is via semantic representation. We could suppose that semantic interpretation is a mapping off of the output of the transformational rules; a supposition more consistent with the current assumptions in the syntax. We could also, for the purposes of the procedure, assume the learner has acquired the base rules for the most part. Then, where there are systematic differences,[16] the learner could hypothesize a transformation.

Second, the discrepancies between the learning theoretic motivation for the constraints and the descriptive (= syntactic) justification of them don't vitiate the learnability work, rather they pinpoint more precisely the areas that require work. Williams' (1980) revision of the Raising Principle takes steps in this direction. Another endeavor in this direction would be to test the syntactic constraints in a theory of learnability; to see to what extent they are sufficient to prove learnability, and in what ways they are inadequate. The distinction along these lines between the descriptive and learning theoretic statements of the Binary Principle discussed above (with reference to Culicover and Wexler 1977, fn. 54) is most interesting

because it involves the question of the cyclicity of S and $\bar{S}$ with specific reference to the Binary Principle. Just this question is raised with respect to Subjacency in the syntactic domain. Such correlations deserve more attention, and should be considered among the potentially positive results of the learnability program at the very least because they point to precise questions that need to be asked.

1.4.3. Another View of the Relationship between Acquisition and Linguistic Theory: The Projection Problem

The work on learnability, which we placed in the second category is, clearly, removed from actual child language data--in which sense it contrasts starkly with the work we discussed in the first category. It is also, as we have seen, distinct from descriptive syntactic work. This characteristic separates it from the work we have placed in the third category. To my knowledge the approach here has been suggested only in the work of Baker (1977, 1978, 1979). It is related to the learnability work in an interesting way. Baker (1977) noted that Culicover and Wexler (1977) themselves saw that the Hamburger and Wexler (1975) learnability proof referred only to obligatory transformations and that optional transformations posed a serious problem for learnability. Baker attributes the problem to the fact that learning them requires negative data: data about which sentences are ungrammatical. One essential problem with optional rules is that many of them have exceptions and require, hence, negative data that is not assumed to be available in the learnability proof. This assumption in fact is correct because such data have been observed to be unavailable to children in the course of acquisition. A second problem for the learnability of optional rules is that a child may, in principle,

have to wait an infinitely long time before being exposed to the relevant datum that would cause the optional rule to be incorporated into his grammar. It is to such problems that Baker's work is addressed.

More precisely, Baker addresses his work to the problem of defining rules in a transformational grammar which have a reasonable chance of being acquired. He sets out the problem with the following claim:

> Classical transformational theory makes available for the description of primary data from English a number of optional transformations. In many cases, these generalizations prove to be false but their falsehood is not apparent until we are provided with the information that certain specific sentences are ungrammatical. This is just the sort of information to which children learning English appear to have no dependable access. (p. 546)

Thus, Baker approaches the problem of learnability from a slightly distinct point of view, and ties it in a unique way both to linguistic theory and to observations from language acquisition. The theory, which defines universal grammar, should, he argues, be restricted in such a way as to prohibit analyses which crucially involve general syntactic rules deducible from positive data, but requiring exception features motivated by negative data. As examples of such rules, he cites dative movement, <u>to be</u> deletion, and the putative optional transformation that would relate the sentences in (16a) and (b); deriving (b) from a structure underlying (a).

(16) a. It is likely that John will be held up.
b. John is likely to be held up.

Baker points out that a theory permitting such a rule would result in a grammar in which (17b) would be similarly related to (17a); an invalid prediction.

(17) a. It is probable that John will be held up.
b. *John is probable to be held up.

The placement of such restrictions on such a rule to avoid predicting (17b) as being part of the language would require precisely the negative datum that (17b) is not a grammatical sentence: a datum of the sort that is unavailable to the child.

To deal with such problems, Baker proposes the following as an outline of a research strategy:

> The first step is to find instances in which a current theory of universal grammar allows simple rules which are compatible with positive data from a language, but which prove to be descriptively inadequate when tested against the full range of adult judgments. The second step is to devise one or more general constraints which would have the effect of excluding these troublesome grammatical rules in advance. The third step is to delineate as carefully as possible the other descriptive consequences of each proposed restriction, to see whether it can be maintained without adverse effects, or whether some other restrictive hypothesis would be preferable. (p. 559)

These steps are proposed to move in the direction of a solution to what Baker outlines in the passage we first cited, and which he names <u>the projection problem</u>:

> Why does a fluent speaker of English consistently make the judgement X, when the reverse judgment would not have been ruled out either by his primary data alone or by a standard transformational theory applied to the primary data? (p. 574)

It isn't hard to see how this problem can be stated explicitly as an acquisition puzzle: How does a child manage to acquire some adult intuition ("judgment X") when the linguistic theory that defines his hypothesis mechanism countenances (even, in some cases, favors) the construction of a grammar, given available data, which predicts the reverse of the intuition (i.e., that some string is not a grammatical

sentence). In the context of this problem in language acquisition, Baker's work becomes even more exciting.

Further, there are two important questions raised by two of the assumptions Baker makes in outlining the projection problem. The first assumption deals with the definition of primary data. Baker does not give an explicit definition of what such data include, but by example shows that they exclude non-positive data and hence, are essentially textual. Thus, information about what strings are ungrammatical sentences, being non-positive data, is not in the set of primary data. Baker also excludes from primary data, information about meaning. As well, he excludes information about ambiguity, synonomy and non-synonomy, and anomaly from the range of primary data. Baker's assumptions about the inappropriateness of negative data are, we have seen, consistent with observations in child language. The logical question to ask is whether the assumptions about meaning are appropriate as well. Is such information available to the learner?

A plausibility argument can be made that children do have access to certain aspects of meaning. Brown (1970) systematically observed that approval or disapproval of children's utterances by parents are linked not to the grammatical form of the utterance, but to the truth value of the proposition. If we assume that children understand these displays of approval and disapproval, and understand that they are indeed directed to a particular aspect of meaning, then children do seem to have access to certain features of meaning.

Recall that the learner's access to the semantic interpretation of a sentence was crucial to the learnability proof in the work of

Wexler and Culicover (1980) and Wexler, Culicover and Hamburger (1975). Whether we accept the assumption that semantic interpretation makes the base available to the learner, it still plays a central role in the construction of hypotheses about transformational rules. It can do so in the way to which we have already alluded in the discussion of that work: it gives the learner access to certain aspects of surface structure (relating to the interpretation of anaphors, including "gaps," for example) and, to the extent that such structures systematically relate (in ways that must be defined) to structures generated by the base rules, the learner can hypothesize a transformation. Restrictions on the form(ulation) of transformations could define the "systematic relationship" necessary to allow formulation of a transformation. Restrictions on the functioning of transformational rules continue to have independent motivation, as Baker points out.

It is difficult at this point to assess the degree of accessability to aspects of meaning that a learner may have. It may be limited to gross aspects of synonomy relating to truth value but it is likely that it does include information necessary for parsing. It is also just as likely that a portion of this semantic information is inaccurate--is not consistent with the actual interpretation that an adult grammar would derive from a given surface structure. This is because the domain of semantic interpretation is the most likely area within which pragmatic considerations will play a role. We have already seen an example of this interplay in children's interpretations of passives (cf. Section 1.3.2 above). In such cases, it may be that invalid semantic interpretation of certain surface structures

but not other identical ones directs children's attention to the structures, and hence provides evidence for reanalysis of the grammar. Such possibilities are intriguing, and though for the moment they remain in the realm of speculation, they can lead to the formulation of specific questions that may be directed toward the data of child language.

If we take information about the semantic relationship synonomy to be within the range of primary data available for grammar construction, what about other aspects of semantic interpretation? Baker is surely correct (note 35) in pointing out that given the present shape of syntactic theory, the fact that a syntactic rule appears to result in an anomalous string does not suggest that the rule must be restricted. The status of interpretation in the theory also makes the role of ambiguity in evaluating the output of transformational rules questionable. Hence, while the question of whether such information is available to the child remains an interesting one, it becomes less crucial in a direct way for the evaluation of syntactic rules in the course of grammatical development.

The second assumption Baker makes that we will treat is really the indirect consequence of an observation. The assumption is roughly that it is easy to tell whether certain syntactic errors reflect the overgeneralization of a transformational rule or not. Baker questions whether children go through a stage during which they commit errors of overgenerality. Such evidence of course would help us determine whether instead the child starts out with the apparently optimal but ultimately overly general rule, to later restrict it. Baker proposes

that in the case of examples such as dative movement, children make only what may be construed as random subcategorization errors. Such errors might in fact look like the sentences of (18).

(18) a. *George said Maxine something terrible.
b. *We reported the police the accident.

The sentences in (18) are the examples Baker cites of ungrammatical sentences that would result from the overapplication of dative movement. It is difficult to determine whether these errors are in fact the result of the overinclusive application of dative movement, or given a theory where the grammar would not have a rule such as dative movement, faulty knowledge of the subcategorization for the verbs say and report. Baker distinguishes errors such as (18) from those in (19), allowing that the latter likely are instances of overgeneralization; in this case, that particle movement and dative movement are not restricted to NPs that are [-PRO].[17]

(19) a. I turned off the light
b. *I turned off it.
c. I gave Sally the toy.
d. *I gave Sally it.

While the point must be made that the distinction is not so clearcut between what may be errors of transformational overgeneralization and errors of subcategorization, it is, I think, correct to propose that such processes as dative alternation pose special problems for acquisition as transformations and likely need to be ruled out as candidates for transformational analysis. This is especially true given the properties they share with other similar phenomena, such as sensitivity to specific lexical items. The point is that this conclusion is not likely to be supported directly by evidence from child language,

because the evidence is hard to evaluate. It's a conclusion that will come from the theory and will have to wait for explicit support in child language until we have a principled way of distinguishing among certain error types. This latter is a goal that should be pursued.

1.4.4. Summary

The work we have discussed here, in 1.4, has proposed the most interesting questions connecting linguistic theory and observation in language acquisition. It has also made clear how difficult it is to deal in these two areas--which should, in principle, be closely connected. The dilemma presents a methodological analog to the Heisenberg Uncertainty Principle. The more closely we observe child language, the further we seem to be from learning about the theory--and hence its role in acquisition. The more we concentrate on the development of the theory, the more we are accused of being ignorant of "what actually goes on in the course of acquisition." The work in learnability is, it seems, an attempt to achieve methodological purity in a sense, since it separates itself both from observation in child language and from the descriptive foundation of linguistic theory. But it doesn't seem to have resolved the paradox.

Nevertheless, the parallel we have drawn between the dilemma facing research in linguistic theory and language acquisition and the Heisenberg Uncertainty Principle is limited. Unlike the situation in Quantum Theory, the paradox here does not seem to be one of principle, but indeed one of methodology. We need to find a way of dealing with child language without obscuring the progress we have made in linguistic theory; of asking questions about language growth that come from

the theory, and we need to find a way of looking for promising questions from the theory to ask. In the discussion of Baker's work, it was possible on the one hand to point out that there are places in which systematic observation of what goes on in child language might have a bearing on what are taken as assumptions in the theory: the case of whether information about synonomy or non-synonomy should be included in the set of positive primary data available to the learner. On the other hand, it was as important to point out where observation from child language may not be so dependable, or provide such a direct test of decisions in the theory, as in the case of the dative alternation errors. The remainder of this dissertation is an attempt to ask a few questions, based on the theory, about language acquisition in such a way that the answers we get from child language can help direct decision making in the theory.

1.5. Assumptions and Claims Fundamental to This Work

In this section we will outline and try to make explicit the assumptions that form the basis of this dissertation, some of which we have alluded to in the discussion of other work in language acquisition. The general justification for work of this type is that claims about explanatory adequacy in linguistic theory are based largely on claims about acquisition. Linguistic theory therefore provides the natural frame of reference from which to ask questions about acquisition, and draw the best inferences possible from the observable phenomena.

Such a justification assumes, of course, the correctness of a theory-bound approach to questions in language acquisition. While we cannot establish, a priori, "correctness" of an approach in terms of the answers it may allow us to get, we can speak of "correctness" in terms of the productivity of a research strategy. We have seen that approaches to child language that deny the role of linguistic theory in the make-up of some sort of (inherent) hypothesis generator have not gone very far toward solving the puzzles of language acquisition, because they have not been able to pose questions about language acquisition. They have generally held that it is correct to pose questions about the acquisition of cognitive constructs, of which language is one. Asking questions that come from language, it seems we will certainly be in a better position to determine the appropriate role of the linguistic mechanism in language acquisition and maybe, have clues about the adjacent role of other cognitive mechanisms in the process. At least we may be able to establish boundaries. Unless we work from a well-defined theoretical framework, it doesn't seem that any of these goals will be in reach. This work assumes, therefore, the existence of some sort of hypothesis mechanism that has, as a substantial component, the principles of Universal Grammar which, along with data from the linguistic community, will help it construct a grammar by limiting the set of hypotheses from which it may choose. It is assumed, for the purposes of this work, that the principles of Universal Grammar include constraints that have, in the course of the evolution of the theory, been found to constrain the functioning of transformations (Ross 1969; Schwartz 1972; Chomsky 1973) or have been

interpreted as conditions on the output of transformational rules (Chomsky 1977, 1980a).

We will now look at the basic claims being made in this work, and the questions about them that will be asked. A fairly well-accepted claim made by workers in language acquisition who accept the theory-bound approach is that children's grammars do not violate the constraints of UG. For the most part, Roeper and others would accept such a claim: he himself has asserted it. We have seen as well, that where the conclusion has been that child grammars appear to be violating some constraint (Matthei 1978), or obeying some constraint foreign to UG (Tavakolian 1978; Solan and Roeper 1978), the interpretation of the evidence supporting the intermediate grammar proposed in these studies must be reevaluated. Further, in the learnability work and in Baker's work, the claim that constraints established are not violated by "immature grammars" is implicit, since the constraints are established as necessary for grammar construction.

Against this claim, we place the observation that children's utterances deviate from those which adult intuitions would establish as grammatical. For the moment we will deal with utterances, since there has been little systematic attempt made to establish and study children's intuitions.[18] Because of all the dangers inherent in paying exclusive attention to child utterances (or child comprehension) we will propose the following as a working assumption. Systematic deviations (errors) that affect the configurational properties of a sentence (i.e., excluding deviations of length of utterances) will be

taken to reflect a particular grammatical system.[19]

We have already seen, further, that no linguistic theory provides the specification of the correct grammar for the hypothesis mechanism. That is, if we return to (4) (even given the fact that there is no longer an infinite set) we see that the hypothesis mechanism may select from (iii) a grammar that assigns structural descriptions compatible with the primary data, but which is not descriptively adequate insofar as it turns out to assign the wrong structural description to other strings. This is essentially the problem Baker sees for linguistic theory. He argues that the theory should limit available descriptions so that a descriptively adequate grammar can be deduced from the available primary data. However, what can be done with the systematic errors we do find in child language and on the basis of which it seems justified to infer grammars for children?

It is just this question that is central to this dissertation. The grammatical systems we infer on the basis of observed systematic errors are misformulated grammars. The misformulations must, based on the assumptions and claims we have outlined, reflect the impact of the interaction of principles of UG (in particular ways) with the data available to the child. Given that there is only a relatively small set of grammars available, there should be a limit to the types of errors we should find. Thus, on the basis of an analysis of some set of observed errors within the framework provided by UG, we may be able to predict other error <u>types</u> that should be able to occur, and predict types that should be excluded. In the best case, we may be able to predict particular errors within the type. These are the goals set

forth here.[20] The following chapters specifically deal with a proposed analysis of the misformulation of Subject-Auxiliary Inversion in child language. In Chapter 3, the problem for the learner of reanalyzing this misformulation is addressed, again within the boundaries of UG. Chapter 4 discusses one other misformulation, dealing with particle constructions in English. In Chapter 5, we look again at the claims made in the course of the study, and consider their extension to findings concerning related problems in the development of grammars of German and French.

Notes to Chapter 1

[1] This is a section in which the authors are discussing Brown and Hanlon's (1970) work with the derivational theory of complexity in child language. Here they propose that the number of transformations involved in the "derivation" of sentences accounts for the relative time of appearance of the sentences in child speech. Such work also seems to be based on the assumption we have been discussing.

Nevertheless, the use of the word theory in the cited section does, most likely, refer to the (notion of) transformational theory and not just the derivational theory of complexity, as both are involved in the context of this discussion.

[2] Braine acknowledges that the system he proposes raises the question of how children get from such a system to a grammar adequate to account for adult competence. He gives three reasons for claiming this is only an apparent objection. First he claims that his model does not necessarily imply sharp discontinuity, rather it allows for gradual learning of transformational grammar. Second he suggests that in fact the apparent generality of adult competence may be due to the buildup of many particular rules rather than to the construction of very general ones. This last point is an empirical one, he asserts, and he claims that it is a feature of his model that it makes it possible to ask this question. His third reason is that his model also makes clear a need for a theory of acquisition of part of speech categories which he maintains are not universal. In his view only semantic notions are universal.

[3] That is, insofar as a set of behaviors can correspond to a model that is not defined by behavior.

[4] There is an interesting sidelight to this misinterpretation. In a note, Derwing cites Fodor from the Miller/Smith volume as well. Specifically he cites Fodor's speculation that a child could be born merely with the innate capacity to learn learning principles, and that the learned learning principles drive acquisition. Derwing proposes that his own alternative proposal for a model of language acquisition develops along such lines. What he does not acknowledge is that Fodor followed his speculation in the cited paper with the following:

> I do not think that the suggestion is true in any significant sense, but that is beside the point. What is important is that the task of characterizing the information the child brings to language-learning is, at least in principle, distinguishable from the question of whether the child's intrinsic information is innate. (p. 106)

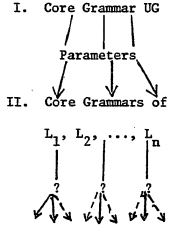
Again, it seems that Derwing bases support for his position on a misinterpretation; in this case, of the thrust of Fodor's speculation.

[5] There is a very definite difference in the way the system is constrained from this point of view now, and the way its constraints were provided for in (4). Note that 4(v), the provision for an evaluation measure, limits the "presumably infinitely many" hypotheses allowed by 4(iii). There is still a provision for an evaluation measure that selects from a class of possible hypotheses for grammars, but that set is not an infinite set. It is limited itself by the various constraints imposed at the levels to which we have alluded.

> To attain explanatory adequacy the theory T must be sufficiently restricted so that relatively few grammars are available, given a reasonable amount of experience E, to be submitted to evaluation; otherwise, the burden on the evaluation procedures is intolerable. (Chomsky and Lasnik 1977)

The role of the evaluation procedure is thus changed: It no longer is the factor that limits the set of hypotheses exclusively. "The theory (UG) outlines a quite small set of possible systems just rich enough to achieve descriptive adequacy"--Core Grammar. Artifacts which represent what may actually be in the minds of members of a linguistic community may deviate from the idealization that is core grammar in (severely) limited ways. These deviations will be characterized by some sort of evaluation system that assigns relative "markedness" to them. Thus is the role of evaluation changed, and the burden of constraining the system shifted (cf. Chomsky 1977, "Conditions on Rules of Grammar," and Chomsky 1979, "Markedness and Core Grammar").

I think this can be made more precise. There is a missing step; shown here as the step from:



where a theory of markedness would be relevant

limited somehow

III. "Actual" mental representations--actual grammars of L_1, L_2, ..., L_n

(class discussion of Pisa Lectures [Chomsky 1980], Spring 1980, led by M. Rochemont)

[6] Roeper (1978), a paper exciting for its approach to the view of language acquisition, separates these 'notions' into two theories of acquisition: the 'pragmatic theory' and the 'semantic theory.' I agree in principle that the notions should be distinct. I do not agree that their treatment in the literature allows us to characterize them as discrete theories; I think that this confusion in the acquisition literature is, further, important to recognize. It reflects particular views of language that we must avoid adopting "by default."

[7] Implicit in my use of the word pragmatics are really two aspects of communication that have been kept distinct by many of those who work on them. The two are pragmatics and functionalism. I have not separated these because the discussion, even stated as it is, making reference to the one, is valid when extended to the assumptions underlying both.

[8] While this is a correct view of how the notion of semantics is involved in this 'language use' view of the acquisition model, I do not wish to give the impression that semantics is not involved in acquisition or that a complete picture of acquisition can be drawn without an account of the development of such knowledge and its relationship to the syntax. There is in (4) 'room' for semantic considerations--interpretation of anaphors, for example, interact with the syntax. But (4) does not allow a view in which semantic notions of this or any sort of control the hypotheses about structure.

[9] In an attempt to control for this fact, Matthei varied simple plural with conjoined NPs in matrix and embedded strings. The precise effect of this variation on the results is not reported.

[10] This latter was suggested to me as a possibility by S. R. Anderson. Thus, each other may have a more secure interpretation in child language with a conjoined antecedent, but a plural reflexive, for example, may not require such a condition. On the other hand, both each other and themselves may show interpretation distinct from anaphors such as them in (i):

(i) a. The chickens said that the pigs were hitting them.
b. The boys want the girls to like them.
c. The dogs noticed the cats chasing them.

[11] Figure 11 is misleading in this regard since it shows any S as a recursive node, not just the initial S, although it is the structure commonly attributed to the hypothesis. To my knowledge, no investigation of multiple embedded structures has been undertaken: that is, we do not know how children would interpret strings like (i).

(i) The sheep that the lion that the tiger saw
_____ _____ hit the cow.

[12] The question of what category defines the clause is one we will deal with in detail in Chapter 3. There is proposed that the category M^{max}, dominating AUX, is equivalent to S.

[13] Roeper observes that structures with the VP 'put NP PP' block the interpretation that he identifies as reflective of the S-initial hypothesis. He concluded that children could therefore know that some S can occur as VP complements, not just attached to S (pp. 120-21).

[14] This principle has a form and consequences reminiscent of certain $\bar{X}$ proposals which suggest universals of constituent structure, but leave as language particular, within specified limits, the aspects of constituent order.

[15] These problems are addressed in recent work. For discussion, see Wexler (1981a, 1981b).

[16] Given a precise notion of systematic, which will also help to restrict the possible hypotheses for transformations (cf. note 19).

[17] The restriction on particle movement, and on the process that results in the dative alternation seems fairly simple and well motivated. There does seem to be support for distinguishing PRO forms of syntactic categories from fully specified lexical forms. Such a restriction, in the case of (20) for example, might be learned on the basis of exclusively positive evidence. The evidence would be the clitic status of these object pronouns, and thus, their (obligatory) attachment to the verb. Baker himself points out that the difference in a stage with (19) and later one that excludes (19b) and (19d) is that the later stage includes "a rule that states that object pronouns are enclitic" (p. 570).

[18] There is only one, informal attempt that I know of, with any success, and that is Gleitman and Shipley (Cognition 1972). See also the discussion of Fischer (1971) in Chapter 4, below.

[19] Systematic has, actually, a two-part definition. The two parts represent two steps in the investigation, and two dimensions to the definition. The definition begins as "non-random," within the boundaries outlined. In the fruitful cases, a grammar inferred on the basis of an error so judged systematic will predict other errors that do in fact occur (in speech, or in comprehension). This clustering of errors represents the second dimension of the definition of systematic. It is the goal of this investigation to achieve this definition, because it will confirm our original judgment and hence, in an important way, the direction of the work.

[20] The most current work in syntactic theory supports this line of reasoning. Chomsky (1979, 1980, but also as early as 1973) and

Chomsky and Lasnik (1977) outline the possibility of parameters along which specific grammars (Core Grammars of particular languages) may differ. There are few parameters, and they themselves are subject to controversy as to the domains over which they may range (Rizzi 1976, 1977; Reinhart 1979), and how they must be limited. This dissertation accepts implicitly the notion of parameters, but does not consistently attempt to test particular ones, partly because of the instability of their individual correctness, and partly because it is not clear how they might be directly tested in the realm of language acquisition. But cf. the discussion in Chapter 3, section 3.2 concerning one such parameter.

Chapter 2

THE PARADIGM CASE

2.0. Introductory Remarks

We will now turn to what serves as the paradigm case for the purposes of this work. This is the account of the appearance of Subject-Auxiliary inversion (henceforth SAI) in the WH and YES-NO questions of child language. To begin, a grammatical analysis will be proposed that will account, in a non-ad hoc way, for the systematic deviations that occur in these constructions. The claim is that this analysis reflects the misformulated child grammar generated by the hypothesis mechanism, and the goals here derive from this claim. First, I propose to show how particular principles of UG are involved in leading to the misformulations in question. These same principles will rule out other logically possible grammars. Thus, the proposed analysis will provide not only a general account of the systematic errors which do occur, but will also rule out other (logically), possible errors. The (systematic) absence of these errors from the data will lend support to this account, and in turn, we will have provided an explanation for their absence: the grammars that would generate them are not themselves generated by the hypothesis mechanism.

The chapter is organized as follows. We will first outline the data that define the problem. Second, we will briefly consider four

accounts from the literature dealing with aspects of the data. The third step will be to outline and motivate the analysis of the misformulated grammar, and to consider the issues raised by the proposal for the misformulated grammars.

2.1. The Observation and the Problem

The data initially crucial to our discussion appear in (1) and (2):[1]

(1) a. Does the kitty stand up?
b. Is Mommy talking to Robin's Grandmother?
c. Did I see that in my book?
d. Will you help me?
e. Can't you work this thing?
f. Can't it be a bigger truck?

(2) a. Where my spoon goed?
b. What he can ride in?
c. Sue, what you have in your mouth?
d. Why he don't know how to pretend?
e. Why kitty can't stand up?
f. Which way they should go?
g. How he can be a doctor?
h. How they can't talk?

The striking fact about (1) and (2), which appear at the same time in children's language, is that while YES-NO questions (1a-f) exhibit the inverted auxiliary, WH- (or, constituent) questions do not. This observation was first made by Ursula Bellugi (1967) and has since posed an interesting acquisition problem.

This particular period is preceded by two other stages in the development of questions. One is quite well documented, while the other is only sporadically indicated by data in various places in the literature. In the well-documented stage (Klima and Bellugi 1973), YES-NO questions are signalled exclusively by intonation, and WH

questions differ from non-interrogatives only in having a WH element in sentence-initial position.[2] There is no inversion in either question type. The other stage involves an apparently brief period during which there seems to be inversion in both types of questions. An example of this phenomenon appears in (6) below. The interesting problem raised by these two stages is that the second one we discussed precedes the period during which there is no inversion. We will discuss the implications of this sequence of events in greater detail in Chapter 3.

2.2. Other Approaches to the Problem

In the following sections are brief discussions of four proposals for an explanation of these data. The four we will discuss are first, Klima and Bellugi (1973; second, Brown (1968) and Brown and Hanlon (1970); third, Mayer, Erreich and Valian (1978); and fourth, Labov and Labov (1978).

2.2.1. Klima and Bellugi

The first extensive study of the development of questions was carried out in conjunction with the development of negation by Ursula Bellugi (1967).[3] It was here that she first observed the phenomenon revealed in (1) and (2). In Klima and Bellugi (1966, 1973) questions and negation are considered and for each, three stages of development are suggested. We have already alluded to the stages for questions in the discussion of the data in (1) and (2) which are, recall, stage three in the Klima-Bellugi account.

Largely, the paper is an attempt to schematize regularities in child speech using the formalisms current in transformational theory at the time. In this way, the paper fits quite well into the research paradigm of the period, which we outlined in Chapter 1 (cf. section 1.0). Hence, the result of the paper is a miniature grammar that includes phrase structure rules and transformations that will generate the children's observed utterances. The question of how the language acquisition device might produce such a grammar is not asked because, as we saw, such questions were not suggested by the goals. We will, in the discussion of our proposed account of a grammar that would generate such data as those in (1) and (2), refer to certain of the rules postulated in this study. The goal, however, will be to consider precisely the question of how such rules themselves might be generated by the hypothesis mechanism that is part of the language acquisition device.

2.2.2. Brown

Brown (1968) too fits well into the research program of its time.[4] It highlights yet another aspect of the framework discussed in Chapter 1. In this paper, Brown essentially treats only WH-questions, and proposes an account of the absence of SAI (which he refers to as "transposing") in them. He first separates questions such as those in (3) from those in (4)

(3) a. What you want?
    b. How you open it?
    c. What his name?

(4) a. What he wants?
    b. What he will want?

c. How he opened it?
d. Why you can't open it?
e. What his name is?

Brown distinguishes the questions in (3) and (4), arguing that questions such as those in (4) are created by a grammatical system that takes echo question ("occasional" questions in the paper) and preposes the WH element. Questions such as those in (3) are claimed not to provide strong evidence for such a source, because they do not contain auxiliaries, be, or any overt verb morphology. He classifies them, therefore, as "weak preposing," and suggests they are in fact derived by "telegraphic reduction" from full questions such as What do you want? How do you open it? and What is his name? (cf. p. 285). The questions in (4) he claims further are derived by preposing, because there is no adult model from which they could be derived by the process of "telegraphic reduction." This particular line of reasoning is based on an assumption that has no real support from grammatical theory, but is nonetheless fairly common, and reminiscent of the phonology study by Neilson Smith (cf. Ch. 1 above, Section 1.1.2.) in which the identical assumption is made explicit; that is, that some adult speech model is the "deep" (or "underlying") structure from which child utterances are derived. In any case, it is difficult to establish the contrast between (3) and (4) given (3a) and (3b) where the echo question would contain no auxiliary or overt verb morphology and thus the absence of these in (3a) and (3b) could not be taken as evidence that these examples were so derived.

Brown makes yet another distinction between such WH-questions as (4) and "why questions." The latter, he claims, cannot be derived by

a WH constituent preposing transformation because the why corresponds to no constituent in a way parallel to the other questions. This claim is based on his observation of children's semantically inappropriate responses to why questions (they do not answer with explanations), and to a view of the adult grammar in which why questions were not derived in the same way as other WH questions.[5] Brown therefore asserts that the why questions (which were most prevalent in one child in the study) are formed by placing why in front of a declarative. We thus have three distinct derivational histories of questions: for forms in (3) and (4) and for why questions, respectively, the latter two being forms occurring in child language at the same time, and having strikingly similar forms.

The explanations for the absence of SAI in these forms is similarly disjoint. In the case of the forms in (4) Brown does not offer an explanation of why only the WH form is preposed, and why there is no SAI. For why questions, the explanation is that why is merely appended to declarative sentences. Brown rhetorically asks why a child's imitation of a correctly formed why question does not exhibit SAI, and answers (for the case of the particular child, Adam):

> Probably he copied according to his present understanding--as children do when they pretend to drive a car or read a newspaper. Perhaps his imitation took the form it did because that form was close to the general operation that Adam was using with his other WH questions. (p. 287)

In essence, then, we have no explanation of why WH-questions should not involve SAI in child language.

Brown and Hanlon (1970) suggest that derivationally simpler (in

terms of the number of transformations involved) constructions may precede more complex ones in the order of acquisition. They apply this suggestion to the case of WH-questions, proposing that

> It is even conceivable that children say What he wants? and Why you went, because they are trying to use a single rule for WH-questions and embedded WH clauses as in: We know what he wants and We know why you went. (p. 203)

There are at least two problems with this conception, the question of derivational complexity notwithstanding. First of all, this proposal allows one to suggest that when SAI appears in WH-questions it might likely be generalized as well to embedded WH-clauses; embedded questions, as above; or relative clauses.[6] Such an extension never occurs, and we are thus left without an explanation of its non-occurrence.[7] Second, one child would use a single rule for YES-NO questions; hence what we would see are WH-questions such as (5)

(5) a. Is his name what?
b. He opened it how?
c. Can he be a doctor how?

In fact we see no such sentences (cf. Dale 1976). Thus, we are still left with the puzzle provided by (1) and (2), having found so far no satisfactory account, and, in fact, no interest in the central question that the data raise: What sort of mechanism would necessarily produce a grammar that generates such structures?

2.2.3. Mayer, Erreich and Valian

Although their work primarily treats the behavior of the auxiliary in YES-NO questions, Mayer, Erreich and Valian (1978) is of dual

importance to the study here. First, they do consider the question of what might shape the syntactic rules of a developing grammar in their account. Second, the account itself raises interesting questions about the development of the auxiliary, which we must consider in connection with any proposals relevant to (1) and (2).

Mayer, Erreich and Valian propose what they refer to as the Basic Operations Hypothesis as a constraint on the formation of transformational rules by the acquisition device. This hypothesis states that for any transformation that combines more than one basic operation (where basic operation is defined as copying, deletion, or insertion) can be misformulated as one of the operations. Thus, a movement transformation, which involves copying and deletion, could be misformulated as a copying rule. A movement could also be misformulated, in this sense, as a deletion, although no such errors are cited or discussed.

Mayer, Erreich and Valian propose that the rules of SAI and tense hopping are misformulated as copying rules. Such misformulations would predict the errors we see in (6) and (7).

(6) a. Did you came home?
b. What did you bought?
c. What shall we shall have?

(7) a. I did broke it.
b. Jenni did left with Daddy.
c. I did rode my bike.

and the following errors as well (from Prideaux 1976)

(8) a. I didn't got...
b. You didn't had some...
c. She didn't goed...
d. The plant didn't cried...

A similar copying rule was first proposed in Hurford (1975) for data

such as those in (6). The important distinction between the account in Hurford and that in Mayer, Erreich and Valian is that Hurford suggests no reason for why the child might develop a grammar with such a rule [this is in fact a criticism levelled against Hurford by Fay (1977, p. 145)], while Mayer, Erreich and Valian do precisely that. Further confirmation for this copying analysis comes from such errors as those in (9) [and (3c)].

(9) a. They didn't spilled.
b. 'Cept you didn't started it, so I started it.

In such strings we see regular verbs involved in the phenomenon as well as just the irregular verbs such as <u>come</u>, <u>buy</u>, <u>leave</u>, <u>ride</u>. If only the latter were involved, an argument could be made for an alternative claiming that the child did not analyze these verbs as carrying tense at all.

In general, the Mayer, Erreich, Valian proposal is a highly interesting one. As they themselves point out, nevertheless, attention must be paid to errors which the hypothesis predicts, but which do not occur. One such error is seen in (10).

(10) What did I see what?

WH-fronting is never misformulated as a copying rule. Goodluck and Solan (1979) propose an account of why this should be so; one which preserves the basic operations hypothesis, but limits it to transformations that do not involve essential variables. We will suggest an alternative account for the absence of these errors in Chapter 3, section 3.4.2. More serious is the absence of errors involving the misformulation of movement rules as deletion rules. Nonetheless, this absence does not vitiate the hypothesis. Rather it points to an area

in which the hypothesis itself is too strong--or has not taken into account the restrictions on types of deletion that may be present in UG, and hence rule out in advance logically possible deletion rules.[8] As well as errors that are predicted and do not occur, there are errors that occur which are not predicted by the framework, such as (11).

(11) I don't had a nap.

Errors like this one, in which tense is clearly not copied could, as Mayer, Erreich and Valian suggest, be counterevidence to the hypothesis unless an argument could be made that such errors are in fact not syntactic, but perhaps lexical.

2.2.4. Labov and Labov

In Labov and Labov (1978), the absence of SAI in WH-questions is specifically treated. It is, however, considered from a unique point of view. Labov observes that there is some variation in the appearance of SAI in these questions. It is not wholly absent, appearing then suddenly. Rather, it develops differentially.

Labov asserts that in the study of acquisition we should be able to deal with (systematic) variation.[9] His assertion is appropriate, given that we want to determine what constrains the construction of grammars over time, and given that we may well be dealing with more than one grammar at some point in time: grammars may "compete" as rule formulations are evaluated, by the hypothesis mechanism, against relevant new data.[10] However, Labov is incorrect in his extension of the assertion in at least two ways. First, he contends that linguistic

theory rules out the possibility of accounting for systematic variation. This is not true. He is focusing on the problem of accounting for the variable application of single rules. Linguistic theory _is_ (correctly) constrained to rule out the inclusion of non-grammatical factors affecting the operation of a rule in the structural description of a rule. Nonetheless, variation among grammars is explicitly allowed within the framework of parameters and no linguistic constraint prevents, in principle, a native speaker from "controlling" more than one grammar. Again, however, syntactic theory does not specify any "contexts" in which one or the other grammar might be selected, nor does it provide a mechanism for doing so. Such considerations are not properly within the domain of a syntactic theory.

The second error in this approach is the assumption that variation in child language is the same as (or quite similar to) variation in adult language. There is, however, a major distinction. Developing grammars vary over time as well as possibly across contexts. Adult grammars do not vary over time--simply, they are not "developing." The result of this assumption is Labov's proposal of a single variable rule of SAI for WH-questions in the grammar of the particular child whose language was observed over a period of seventeen months.[11] This rule recognizes explicitly only the variation across types of WH-questions--it reflects, for example, the observation that this child's _why_-questions were least "susceptible" to SAI. However, Labov cites 41 factors in 8 groups that might account for the observed variation.

As a consequence of this account, Labov misses the possible

questions one could pose about the effect of the development of parts of the grammar (AUX, for example) on successive formulations of SAI. For example Labov observes:

> For reasons we do not yet understand, the past tense consistently favors inversion. (p. 20) [He is referring to past tense endings on main verbs]

Presumably this observation includes the appearance of DO in the appropriate syntactic contexts. If we look at the possibilities, it is not hard to conjecture that at some point the category AUX in the child's grammar includes TENSE and DO, but no rules moving other verbs (e.g., have, be, modals) into AUX. So, a rule of SAI moving AUX would move exactly tense and DO. The absence of present in this observation could have a morphological explanation: Past is morphologically marked consistently, while present is not, and hence may be "harder" to pick out. There are examples of correctly inverted DO with mismatched tenses (analogous to (11) above) in the literature, a phenomenon that would follow from such an analysis. While this is just speculative, it does outline particular data that are predicted (the specific types of mismatch) and proposes an account for the phenomenon absent in Labov.

Labov hence is mistaken in his claim that "analysts are primarily interested in describing a stage of development within a homogeneous view of language and discount [underline mine] variations as remnants of previous stages or anticipations of new ones" (p. 3). Variation, as it relates to syntactic theory is very important, particularly over time, since the crucial question in child language, again, is what constrains the changes in grammars. An analyst concerned with such a

question does not discount variation as remnants of overlapping "stages," rather it is precisely these overlaps, and how the grammars that produce them are defined that is interesting.

It seems, then, that the puzzle of (1) and (2) remains. Regardless of the variation in the appearance of SAI, the fact that it does fail to appear, systematically, is still not explained. Why should the rule be blocked in WH-questions, but not in YES-NO questions. In the next section we will propose a possible account for this phenomenon, and consider in detail how the hypothesis mechanism, given principles of UG, would arrive at such a grammar.

2.3. Relevant Fragments of the Developing Grammar

We will now consider what the relevant parts of the grammar that would account for (1) and (2) should look like, and how they might be constructed, given a fairly restricted theory of grammar. We will begin with the phrase structure rules, focusing on the categories COMP and AUX. The first claim is that we are justified in positing COMP as a category in the (child) grammar.[12]

The evidence in support of this claim will come from two sources: support for the independent claim that COMP is a category of UG, and data from child language which are best accounted for by a grammar including COMP.

2.3.1. COMP

Arguments for including COMP as a category of UG can be traced as far back as Bresnan (1970, 1972) who, having provided arguments that WH is a complementizer, generalized Baker's (1970) Q-universal into the COMP-substitution universal:

> Only languages with clause-initial COMP permit a COMP-substitution transformation. (1970, p. 261)

Research has revealed an intricate relationship between the position of COMP in a language and a number of other factors: the order of major constituents, types of question formation rules, and the presence of sentence-initial question particles (cf. Greenberg 1966, for example).

The availability of sentence-initial COMP as a category in UG sets as a question the determination of how it is selected by the hypothesis mechanism as a category in the developing grammar. That is, is there a principle that the mechanism can use to discover whether the language for which it is constructing a grammar exhibits evidence for COMP? Focusing on the potential for an implicational relationship between the factors cited above as well as a number of Greenberg's (1966) Universals relating to these factors, and the presence of sentence-initial COMP, Emonds (1979) proposes the following generalization, the basis for which he attributes to work by Den Besten (1977).

> All instances of movement to a pre-subject position by a grammatical transformation are attractions to a sentence-initial COMP node. (p. 76)

This principle extends Bresnan's universal, suggesting that a wider

class of preposing rules entails the presence of COMP in a language. This extension in turn suggests the following as an implicational principle:

> If a language exhibits evidence of preposing rules, it has a sentence-initial COMP.

However, there is a problem with proposing any of these as a principle that the mechanism makes use of in the course of grammar construction. There is, in fact little evidence of preposing rules early in the developing grammar in cases that would be crucial. WH-questions, for example, provide no evidence that would justify an inference on our part that the grammar contains a rule of WH-movement. In fact, a current research problem in grammatical theory is the determination of evidence for movement rules (cf. Chomsky, "Principles and Parameters" 1980).

Nonetheless, there is a common thread in the three generalizations which, in light of developments in UG subsequent to their proposal, does suggest the outline of a principle which could lead the hypothesis mechanism to posit COMP in the grammar it is constructing. Whether or not there is a movement, all three generalizations suggest that there is a sentence-initial category in which a particular set of grammatical formatives is generated. It is the nature of those formatives that they bind either a variable or a trace[13] elsewhere in the clause, and in order for the binding to be proper, the relation of C-command must exist between the formatives and the elements they bind (for details cf. Chomsky 1973, 1977, 1980). We will assume, then, that COMP is a universal category, and that what the mechanism must deduce, is whether a sentence-initial COMP is correct to posit for the

grammar of the language in question.[14] That the deductive process is fairly straightforward, and that the developing grammar shows evidence of COMP quite early seems to be reflected by the data that are available to us, an examination of which we turn to now.

In their study of the development of questions and negation, to which we referred above, Klima and Bellugi (1973) propose the following phrase structure rules to account for the preliminary stages they observed in the acquisition of questions and negation.

(11) a. i. $S \rightarrow \left\{ \begin{array}{l} \text{Q yes/no} \\ \text{Q what} \\ \text{Q where} \\ \text{Q why} \end{array} \right\} -$ Nucleus[15]

ii. Nucleus $\rightarrow$ NP - V - (NP)

b. $S \rightarrow [\left\{ \begin{array}{l} \text{no} \\ \text{not} \end{array} \right\} -$ Nucleus$]$[16]

The question that should arise in the face of such rules, is why we should see these elements (Q and NEG) precisely in sentence initial position. While someone might propose, in the case of WH questions at least, "because children systematically hear them there," such an explanation is not available in the case of negative sentences. Nor does an explanation based on some notion of perceptual salience hold. That is, one cannot suggest that these elements occur sentence-initially because that is a position of high prominence in a sentence. First, sentence-final position seems an equally prominent position in English--primary stress generally falls here. Further, as noted in Drachman (1978), results in an imitation study done by Blaisdell and Jenson (1970), showed that children recalled best syllables that were

stressed, or that were last in their string, suggesting that 'recency' plays a more significant role than 'primacy' in children's auditory memory. This in fact confirms Slobin's (1973) claim that the end of a word seems to be perceptually salient. So if there is any content to a perceptual salience argument, it is to be found when the argument is made concerning string- or sentence-final elements. But sentence- or string-final position is not the position we find Q and NEG in. The question is, then, why these elements occur sentence-initially. Further, these are not the only elements that appear here. Children have been observed to go through a stage where the utterances of the type in (12a) have the interpretation in (12b).[17]

(12) a. Only I want candy.
b. I (only) want (only) candy.

We have noted (Chapter 1, section 1.4) that Roeper uses facts such as (11) and (12) to support the S-initial hypothesis, and we have seen some objections to such a hypothesis; a major one being that nothing in UG accounts for why the language acquisition device would entertain such a hypothesis about syntactic structure.

In a grammar with COMP, we have an account of why such elements occur in sentence-initial position if we assume that they are placed in the category COMP. The interesting question then becomes, what might justify such an assumption. An answer here is suggested by a combination of the data children have and the principles of UG. In part we return to the account we rejected above as an explanation for the appearance of (11) in the grammars of children: "children generate questions with sentence-initial WH-elements because they hear the elements in this position." The crucial difference is the

interaction of the sentences children hear, and the principles of UG which go into the formation of the grammar. One such principle comes from the suggestion of Chomsky (1977) that WH-words should be considered as quantifiers for the purposes of rules of interpretation which apply at the level of Logical Form (LF). In conjunction with this is the generalization that for a variable to be properly in the scope of a quantifier, the quantifier must C-command (in the sense of Reinhart 1976) the variable (cf. May 1977). Together, these principles insure that quantifier-elements will be generated in a position where they will C-command the variables that they bind. Taken together with the data available to the child--WH-elements occurring sentence-initially, and the assumption that COMP is an available category, these principles can very plausibly result in a grammar in which quantifier-like elements (quantifiers, WH, and negatives) are all generated in COMP. Hence, we have an explanation for the appearance of these elements in sentence-initial position.

2.3.2. The Category AUX

This explanation further predicts, in an interesting way, another feature of the developing grammar. This feature connects negatives and the category AUX. Thus far we have proposed that negative elements are generated in COMP position at the stage which was originally described in (11b). The structure, however, is somewhat more complex. The argument to be made is that certain particles, which seem to be merely negative elements are primitive constituents of AUX at this early stage. Hence, AUX is generated in COMP position as well as particular negative particles. To begin the argument, we review the

sequence of developments observed in child language. Immediately following the sentence-initial placement of NEG elements, exemplified in (13) is the stage that we can see in (14).[18]

(13) a. Not Tom here.
b. No I go to school.
c. No want stand head.
d. No play that.
e. No Fraser drink all tea.

(14) a. You can't dance.
b. Don't leave me.
c. I can't catch you.
d. I don't want it.
e. I no want envelope.
f. I no taste them.

At the stage exemplified by the utterances in (14), there are no utterances such as (15) that have been observed.

(15) He would not give me the crayon.

That is, there are no negative elements that co-occur with AUX elements without being contracted to them. Further, Bellugi observed no independently occurring positive auxiliary elements. This last observation led her to conclude that at the stage where (14) occurred, there is no justification for the category AUX in child grammars, and that the change from (13) to (14) involves only a reanalysis of the placement of NEG.

That the data in (13) and (14) suggest only the progressive analysis of where NEG particles may appear in the sentence is not an implausible conclusion. The alternative claim--that what we are seeing are primitive constituents of the category AUX which is present in the grammar--does not follow readily from these data. Nevertheless, the latter alternative takes steps toward answers to some interesting questions in a way that the former does not. First, if can't and

don't are analyzed in the intermediate grammar as additional members of the set of NEG particles, undifferentiated from no or not, then we would expect them to appear wherever no or not appear, and with the same interpretation as these particles. But they do not. That is, utterances such as (16a) with the interpretation of (16b) do not occur.

(16) a. Don't you dance.
b. NEG you dance.

And, the string in (14b) is interpreted as an imperative, not as a simple negative.

Second, subsequent developments observed in child language raise another question. Consider the examples in (17) and (18) (Kuczaj).

(17) a. He's do take his, take his clean pants off.
b. It's don't have any oil in here.

(18) a. Did you came home?
b. Can you broke those?
c. Don't he wanted to help somebody?
d. Is Georgie wake up?

If we denied the possibility that the elements in question in (13) and (14) were (primitive) constituents of AUX, we would have no account of why, in (17) and (18), positive auxiliary elements do appear fairly soon in precisely these positions, alternating, in the case of do, for example, with the NEG elements. Such mutually exclusive alternation of elements has traditionally provided syntactic evidence for the claim that the alternating elements are members of the same category, and that type of argument should not be ignored for this case.

To summarize, briefly, it has been observed that negative elements in preverbal position appear, in child language, just prior to the appearance of auxiliary elements here, and that the two sets of

elements alternate in preverbal and sentence-initial position. There are three accounts that could be proposed for these observed phenomena. The first is that at an early stage there is no category AUX. There is only a set of NEG elements that first appear sentence-initially and then appear preverbally. AUX is not a category of child grammars, in this account, until non-negative AUX elements appear preverbally. This account largely characterizes the position of Bellugi (1967) and Klima and Bellugi (1973). The account generally leaves the development of AUX, apparently out of the category NEG, as well as the alternation of the two, unexplained. It deserves mention that the claim that elements such as don't are (unanalyzed) NEG, essentially not distinguished from no or not in the same position raises an interesting paradox. We have discussed claims made in the literature, and supported, generally, by observation, that children attend to the ends of words, and that this attention may function as a perceptual aid to hypothesis construction. This claim suggests that analysis (i.e., segmentation) is an integral part of the acquisition process. Therefore, proposing that don't reflects a progression from negative elements to the full range of auxiliary elements provides a more principled step in accounting for the changes from (17) to (18) (as well as subsequent ones) than does asserting it is not an auxiliary, since the former position, but not the latter, would acknowledge the role of segmentation in the process of acquisition, but the latter largely ignores it.

A second, intermediate account allows that AUX is a category in early grammars, as well as COMP, and suggests that NEG may first be

generated as a constituent of COMP, and then is reanalyzed as a constituent of AUX. This account, however, would still leave unexplained the alternation of the two sets of elements, AUX and NEG, in precisely preverbal and sentence-initial position. The third is the account proposed here; that AUX is a category learned quite early, as is COMP, and that what is learned more slowly, reflected in the data, is the constituent membership of the two categories.[19] Hence, the claim is that in the child's grammar these particular "NEG" elements, can't and don't, are constituents of AUX. It is for this reason that AUX elements appear to develop "out of" NEG elements, and that the two sets of elements alternate.

The question does remain why the first AUX constituents are NEG. One answer is that, at least in the case of DO, it is in a negative form that children are most likely to hear (it) in preverbal position. For example;

(19) a. Elizabeth doesn't eat sand.
b. Don't eat sand.

This claim is not, however, based solely on the relative frequency of negative exhortations in the data the child has access to (i.e., (19b)). Rather, it is the case that positive instances of DO in preverbal position--emphatics--are semantically marked, induce intonation changes, and are likely to be fairly rare in the data. In fact, when there are instances of the positive auxiliary DO in preverbal position in child language they turn out not to be examples of emphatics. They appear to be a generalization of the principle that other auxiliary elements can occur in this position; as (17a) showed, and as we

see in (20) (data from Kuczaj).

(20) I did fell when I got blood.

We would expect, then, strings such as (17a) and (20) would appear later than the uniquely negative instances of AUX and, as we have seen, they do. However, we would also expect them to occur in conjunction with the more general appearance of non-negative modals, such as can, for example. As the question seems never to have come up in the literature, we know of no systematic data that bear on the prediction here.

A second question which is raised by the arguments that the negative particles don't and can't are (primitive) constituents of AUX is why they should be so assigned by the LAD. That is, what justification is there for claiming that the initial hypothesis is that these particles are constituents of some category (AUX), rather than just a set of particles that can occur in one of two positions (sentence-initial and pre-verbal)? A corollary to this question is the question of how much evidence is required by the hypothesis mechanism before it assigns some element to a category. Neither of these problems has a straightforward solution. There is, nonetheless, some support for a principle of acquisition that favors category assignment as an initial hypothesis, with relatively little data required for the mechanisms to entertain the hypothesis. The framework within which the grammar develops--UG--gives the notion category a fairly central role in the definition of base rules, the definition of possible transformational rule, and in several of the constraints on the operation and form of transformational rules (e.g., the A over A condition or

Emonds' structure preserving constraint). One can imagine then that the determination of category is a significant first step in the construction of a grammar. Hypotheses in this direction would be, thus, entertained fairly readily, even when they result in misformulated grammars.

Support for the claim that this principle is part of the hypothesis mechanism comes not from close scrutiny of what children do. Nor is the principle itself, as a constraint on hypotheses entertained by the mechanism, the result of some inductive process in the course of language development. In either case the principle is fairly severely underdetermined by the data. It has been shown that whether or not there is (wide) variation across children of data that serve as input to acquisition, the data are not, in any case, "tuned" to acquisition in any significant way (cf. Newport, Gleitman and Gleitman 1977). There is, thus, no way that children could inductively discover a general principle such as "determine category"--or any principle constraining the shape of grammars, since the data required by such a process are not readily available to them. Nor does the set of children's utterances available to us directly reveal the operation of such a principle in the construction of grammars. We can only see its reflection in the intermediate grammars that we infer. What we must say about this principle, and in general of principles that we assume to be part of UG (or of acquisition "strategies" which are derivative of principles that are part of UG) is that it be used abductively[20] by the mechanism in the construction of grammars. The theory of grammar--more properly the research program within which the

theory itself can be constructed--leads us to think in this way, and thus is what ultimately provides the support we require to suggest that such principles do in fact play a central role in the acquisition process, at least with respect to the development of the grammar.

2.3.3. AUX as a Constituent of COMP in the Developing Grammar

We will now consider the status of the category AUX as a constituent of COMP. Given the arguments above that at least can't, don't, in the data above are early constituents of AUX, the generalization that negative elements are quantifiers, and take scope, and the general principle of UG, which we have already cited, that a quantifier must C-command any variable properly interpreted as being within its scope, COMP is highly plausible as a position in which AUX is generated. The claim to be made is, therefore, that there is a stage during which AUX is generated by the base rules as a constituent of COMP as well as a constituent of S in preverbal position.

This claim is confirmed by subsequent observations. At a stage just prior to, and, in some cases, overlapping with the stage reflected by the utterances in (1) and (2), utterances such as those in (6) occur as well as in the following:

(21) a. Did he broke...?
b. Did we brought...?
c. Did we left...?

The strings in (6) have, as we discussed, been claimed to support the hypothesis that intermediate grammars include a rule of auxiliary copying. At the stage where both (6) and (21) occur, however, we could very well be seeing the reflection of the operation of the base

rules, generating AUX in both COMP and preverbal position. Further support for this position, as opposed to the position consistent with the copying analysis, comes from strings such as (18b) and (18c) for which a copying analysis is not possible. Hence, the claim is, that at the stage reflected in (6), (18) and (21), there is a rule in the developing grammar with roughly the following form (ignoring for the moment other constituents of COMP, represented here by "...").

$$\text{(22)} \qquad \text{COMP} \longrightarrow \left\{ \begin{array}{c} \ldots \\ \text{AUX} \\ \ldots \end{array} \right\}$$

And, as the arguments in 2.3.2 propose, there is likely to be a set of lexical statements that assign the lexical items <u>don't</u>, <u>can't</u>, to the category AUX.

2.3.4. "Relating" the Intermediate Grammars

The next step is to investigate the relationship between the grammar that generates the strings in (1) and (2) and the grammar we have been considering, which generates strings such as those in (6), (18) and (19). The questions we need to ask are (i) what are the differences between the latter grammar and the former? and (ii) what might be the mechanism that makes it possible for the former to develop out of the latter? These questions will deal in turn with the problems of determining the point at which we see evidence for transformational rules in intermediate grammars. In the case here, we will consider this question in light of the generation of WH and AUX constituents in these strings.

In the case of AUX, the question involves determining the point

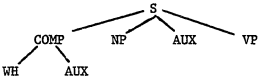
at which the grammar includes a transformation resulting in auxiliary elements appearing to the left of subjects in questions, and determining the form of that rule. There is no evidence that the grammar of which (6), (18), and (21) are reflective includes such a rule. But it appears that the grammar we can infer from (1) and (2) does include such a rule. Added to the development of this rule is another interesting problem. This involves the strings (6b) and (6c), here repeated for ease of exposition.

(6) b. What did you bought?
c. What shall we shall have?

Apparently, the change from a grammar that base generates AUX in COMP to one that moves AUX there also involves a reanalysis of COMP.

Assuming that the strings in (6) do not involve epiphenomena, but do reflect the grammar (i.e., we may infer that the grammar does assign structural descriptions to them), it is possible that the grammar at this point generates structures such as (23).[21]

(23)



That is, the grammar allows a branching COMP. The change from this grammar to the one reflected in (1) and (2) thus, also involves a move away from this analysis. It is this sequence of analyses that we need to look at more carefully. This is because the analyses of AUX and WH in both grammars deviates in significant ways from the generally accepted analyses of these phenomena for developed grammars.

Let us focus initially on COMP. First (as we have noted) the

node COMP may branch; that is, it may be doubly filled. Second, it is not a sister to S, but a daughter (cf. note 14). Both of these features of COMP in the developing grammar result in its being significantly different from the developed grammar posited for English in which COMP is not expanded by the base rules as a branching node, and is a sister to S, not a daughter of S. The change from the grammar without a transformation moving AUX to one including such a rule involves thus a change in the status of COMP as a (doubly) branching node as well as a change in its dominance relation to S. Because these changes accompany the appearance of the transformation moving AUX, it is difficult to determine precisely why the grammar changes without looking at the development of the rule as well.

The claim I am making is that SAI develops initially as a COMP substitution rule. In the wake of this development, COMP is reanalyzed as a non-branching node, in which WH-elements may be generated. There is no evidence in child language of a reanalysis that involves the positing of a rule moving WH at this stage. As we noted in our discussion of Brown (1968), there are no echo questions at this stage. The evidence we have for inferring the change with AUX is fairly straightforward. Auxiliary elements no longer co-occur in sentence-initial and preverbal position.[22] The evidence for inferring that the rule is formulated as a substitution is that AUX and WH do not co-occur in COMP at this stage. The justification for such a formulation is different. That is, the evidence that leads us as analysts, to posit such a rule as present in the grammar is not what we must speculate drives the hypothesis mechanism to construct such a rule. In this

case, the proposal is that formulation of the rule as a substitution is favored over a formulation that specifies an adjunction, which, we will propose in the next chapter, is the form of the developed rule. This proposal is based in part on the assumption that a substitution is simpler because it does not involve the creation of structure.

What makes this proposal for the operation of the mechanism interesting in this case, is the fact that the mechanism "ignores" data in the environment to the extent that it constructs a reanalysis that results in "malformed" WH-questions--ones in which there is no SAI. COMP is no longer branching. AUX inversion is formulated as a substitution, and hence is blocked from applying when COMP is filled with WH. We return here to the question of why there is a restructuring of the grammar away from the ability to generate structures such as (21).

This is not a problem whose solution is straightforward or easy. One possibility is that the strength with which a substitution is favored forces, as an automatic consequence, the reanalysis of the target node as either empty or filled. A second possibility focuses not on the formulation of the rule, but rather on COMP itself. It has been proposed that English has a surface filter ruling out a doubly filled COMP (Chomsky and Lasnik 1977) (cf. the discussion in note 30). Thus, it is possible that this filter is acquired early, and this reanalysis of COMP in turn also favors the initial construction of SAI as a substitution. On the other hand, the content of the filter specifically rules out the co-occurrence of WH and a complementizer in English, accounting for the ill-formedness of strings such as (22).

(22) *John questioned the bloke who that the constable suspected.

The filter does not extend to any constituents other than complementizers. Specifically, it does not refer to auxiliary elements, so its relevance to a reanalysis here is questionable. Another way of ruling out a branching COMP has been proposed by Emonds (1976).[23] He argues that COMP substitutions are root transformations, and only one applies in a given clause. This account rules out such strings as (24).

(24) a. Who never will he support?
b. On that shelf who put the book?
c. These steps what did you never sweep with?

d. *Never who would he support?
e. **Who on that shelf put the book?
f.***What these steps did you never sweep with?

This latter account, involving the COMP substitution principle--if it is part of UG--suggests an interplay between the formulation of SAI as a COMP substitution and the conclusion for the developing grammar, that there is a restriction against a branching COMP. A problem with supposing such an interplay is that subsequent research suggests that facts such as (24), accounted for by the COMP substitution principle (and the claim that movements to COMP other than WH fronting are root transformations), may be the consequence of the interaction of other features of UG (cf. the reference cited in note 23, and Chomsky 1980 [on Binding]). Hence, if it is the case that COMP does not branch (in English, at least), it may well be as the result of other, independent factors; not the result of an explicit stipulation such as the COMP substitution principle.

An important point to note here about the claimed formulation of

SAI in the developing grammar is that there does not seem to be any support for the inference that it is a structure preserving substitution (in the sense of Emonds [1976]).[24] Although in the stage just preceding (the one reflected by (6), (18) and (21), recall), we have argued that the base rules generate AUX in COMP, we cannot conclude that the first formulation of SAI is structure preserving. This is because the structure of UG does not favor such a formulation in this case. For us to assume such a formulation, we need to claim that AUX is generated in COMP as obligatorily unfilled. As Emonds (1976) pointed out, in the analogous case of the problem of WH-fronting,[25] such a requirement would strip the structure-preserving constraint of any content, since we would be claiming that COMP dominates any category in the base, but that the category must be empty, pending the operation of some structure-preserving substitution. In the case of WH-fronting, a weakening of the structure-preserving constraint is proposed in the form of the Sentence Boundary Condition (p. 112):

> The Sentence Boundary Condition: If Aj is the rightmost or leftmost constituent of an S, a transformational operation that substitutes B for Aj is structure-preserving if B dominates A_i, provided that there is no S such that $B = X\ [_S\ Y\ A_i\ Z]\ W$.

Thus, given that COMP may dominate the feature [+WH] in this framework, and the fact that COMP is the leftmost constituent of S, any (major) category dominating the feature [+WH] may be moved, by a structure preserving rule, to COMP. There is no feature, extant, that allows us to suppose a similar analysis for SAI.[26]

2.4 Discussion: Theoretical Issues Raised by the Rules Proposed for the Intermediate Grammars

By way of summary, we will look more closely at the rule systems we have been describing for intermediate grammars, and consider further, issues raised in proposing them. The structure in (23) entails a rule scheme such as (25)

(25) a. S $\longrightarrow$ COMP NP (AUX) VP
b. COMP $\longrightarrow$ (WH) (AUX)

We have already noted that the structures generated by these rules deviate from what has been proposed for the adult grammar of English in at least two major ways. First, COMP is generated as a daughter of S, rather than a sister to it. Second, COMP is branching in this intermediate grammar. In order to make the assumption that a grammar including (25) is possible as a formulation of the hypothesis mechanism, we must claim that the determination of the shape of these grammars is somehow parameterized (in the sense of Chomsky [1980]). That is, the claim is that principles of UG make available to grammars analyses in which COMP either branches or it doesn't, and in which it is either a sister to, or a daughter of S.

In order to argue that such parameters do exist, we need to look at the behavior of rules and possible explanations for such behavior. In the case of the placement of COMP as a potential account there is an interesting possibility for such evidence. A proposed explanation for the apparent violation of Subjacency[27] in languages such as English, which permit WH movement to apply as in (26), is that WH movement does apply successive cyclically; first from S to COMP, and then from COMP to COMP.

(26) Who did John believe Mary thought the Union supported t?

(where t is the most deeply embedded trace left by movement of <u>who</u>). However, even the output of such movement, nonetheless, appears, on the face of it, to violate other constraints on the operation of rules argued to be part of UG; the Specified Subject Condition (SSC) and the Tensed-S condition (or the Propositional Island Condition [PIC]) (cf. Chomsky 1973, 1977; Chomsky and Lasnik 1977). In particular, the PIC as it is stated in Chomsky (1977) states that in a structure such as (27), where α is one of the cyclic nodes or NP, no rule may involve X and Y, where α is a finite clause.

(27) $\ldots X \ldots \underset{\alpha}{[} \ldots Y \ldots] \ldots X \ldots$

This statement results in the restrictions schematized in (28).

(28) $[_{COMP} \ldots] \ldots [_{\bar{S}} \; [_{COMP} \ldots] \; [_{S'} \ldots]]$

That is, even movement from COMP to COMP, which is possible in English, is ruled out by the statement in (27). Added to (27), therefore, is the stipulation in (29) (Chomsky's (46) p. 85, 1977).

(29) Where Y is not in COMP.

In this way, COMP becomes an "escape hatch," but its potential status as such requires a specific stipulation, and (29) is thus, a language particular option. The reformulation of the PIC as part of the opacity condition on the binding of anaphors (cf. Chomsky 1980), which we see in (30) made this stipulation unnecessary.

(30) If α is an anaphor in the domain of the tense . . . of β, β minimal, then α cannot be free in β, where β = NP, $\bar{S}$.[28]

The status of COMP as an escape hatch in a language that permits unbounded WH fronting in this formulation, falls out of its position outside the domain of tense.

As Chomsky (1980) points out, the unbounded character of WH movement, reflected in (26), is language particular, and, as Koster (1978a, b) argues, may be the marked case, in fact. Thus, cases of unbounded WH-movement would have to be stated as exceptional. But neither the PIC (nor its reformulation, the NIC) provide an account of this fact. Chomsky suggests that the matter may relate to the selection of particular nodes as bounding for subjacency; that $\bar{S}$ is universal, and some languages may also choose S (as does English). Thus, WH-movement will always be bounded unless otherwise specified. The specifications become lexical; matrix verbs may be specified such that $\bar{S}$ is not a bounding node for their complements.

However, the other possibility is that there is a difference in the phrase structure rules, such that in languages that do not exhibit unbounded WH-movement, COMP may be generated as a daughter of S, and hence be in the domain of Tense. This is the possibility suggested here.[29] Given that bounded WH-movement is the unmarked case, we would expect that the hypothesis mechanism would initially entertain a grammar including a rule such as (25a). Only the appropriate evidence would cause the grammar to be re-evaluated, and the problem of appropriate evidence in this case is a question we will consider in Chapter 3.

The claim made here as well about the possibility of COMP to branch requires the same line of investigation as we suggest for the determination of whether there is support for the parameterization of the placement of COMP. It turns out that such evidence may exist in Swedish. As Middle English did, Swedish allows WH and the general complementizer, analagous to 'that,' both to appear in COMP.[30]

(31) Jag undrar vem som Maja seglade med.
I wonder who that Maja sailed with.
(Anderson 1975, cited in Baltin 1978)

If languages do differ with respect to whether they permit particles to co-occur in COMP, then we may infer that the explanation of COMP assumed in Chomsky (1980) and Rouveret and Vergnaud (1980) is not universal.

Thus, what we have in the intermediate grammar including (25) is a combination of particular realizations of each of these possibilities that we are claiming are available given the principles of UG. An interesting question arises about what might govern the combination in the developing grammar. To my knowledge, there is no evidence that the placement of COMP and its ability to branch interact significantly in grammars, or that the interaction in any particular way is favored by any principles of UG. Thus, we cannot appeal exclusively to UG for an account of the combination. Nor, however, should we ignore the possibility that some aspect of the theory, as yet unknown, does favor the combination. We cannot ignore such questions just because they are presently unanswerable; they should be kept in mind as variations in the theory developed (cf. the discussion in note 29, above). On the other hand, it is entirely plausible that (25) reflects the

convergence of two separate lines of development, and that there is no correlation between the branching and the placement of COMP.

As a final step in this chapter, let us consider the formulation of the early rule of SAI as a COMP-substitution rule. The claim is that this rule is formulated as (32):

$$(32)\quad [_{\substack{\text{COMP}\\ [+\text{WH}]}}\quad] - \text{X} - \text{AUX} \Longrightarrow 3 - 2 - \emptyset$$

The introduction of this rule entails a grammar including base rules that differ from these in (25) and which supplant them. In particular, (25b) is replaced by (33).

$$(33)\quad \begin{array}{lll} \text{a.} & \text{COMP} & \rightarrow \pm\text{WH} \\ \text{b.} & [+\text{WH}] & \rightarrow \left\{\begin{array}{c}\text{who}\\ \text{what}\\ \text{which}\\ \text{etc.}\end{array}\right\} \end{array}$$

The output of these rules includes the following; where e, by convention, represents an unexpanded category.

$$(34)\quad \text{a.}\quad [_{\substack{\text{COMP}\\ [+\text{WH}]}}\ \text{e}\ ]^{31}$$

$$\text{b.}\quad [_{\substack{\text{COMP}\\ [+\text{WH}]}} \left\{\begin{array}{c}\text{who}\\ \text{what}\\ \text{which}\\ \text{etc.}\end{array}\right\}]$$

$$\text{c.}\quad [_{\substack{\text{COMP}\\ [-\text{WH}]}}\ \text{e}\ ]^{32}$$

Thus, the grammar that we will consider in relation to the developed grammar in Chapter 3 includes (25a), (32) and (33). In the next chapter we will compare these rules with the rules proposed for developed (adult) grammar, preliminary to stating what we will call the

Dissonance Problem and consider what allows the intermediate grammar to be reanalyzed in the direction of the adult grammar; the Learning Problem.

Notes to Chapter 2

[1] These data are taken from Klima and Bellugi (1966, rpt. 1973).

[2] Klima and Bellugi (1973) argue that these very early WH questions only superficially resemble questions in which the object (or more generally, the complement) of the verb has been questioned and preposed. They assert, that is, that children do not understand constituent questions. This assertion is based on their interpretation of the following interchange as exhibiting an inappropriate response from the child:

i.	(a)	Mother:	What, did you hit?
	(b)	Child:	Hit.
ii.		Mother:	What did you do?
		Child:	Head.

However, if the comma in i(a) is correct as it appears in their data, the question becomes in fact a YES-NO question (albeit a marked form) to which the child assents, by repeating what he/she has interpreted as the FOCUS. In (ii) it is possible [given that (i) and (ii) are part of the same discourse] that the child has been given in (iia), the correct interpretation of (i), and is, in fact answering that, using do as an unmarked replacement for the V, hit, rather than as the "marker" for the questioned VP. Note that this may be a place where pragmatics can be viewed interferring with syntax. In the interest of preserving the discourse, the child ignores the details of the question (ii) to construct a response to his revised interpretation of question (i). A further possibility that cannot be ruled out is that Head is mistranscribed, and that the child indeed uttered (some phonological variant of) hit, which the transcriber did not understand. Apparently these data were not all recorded at the time in phonetic transcription (Brown, "The Child's Grammar from I to III," p. 103).

Whatever the correct interpretation of this datum may be, nonetheless, it is not evidence that children do not understand constituent questions at this time. We are finding, more and more, that many studies of this type have underestimated children's comprehension of linguistic structure at various chronological points.

[3] But cf. as well Bellugi (1965).

[4] It should not be surprising, of course, that these studies should reflect the same assumptions because they are so closely related. In fact, they are all part of the larger naturalistic

longitudinal study carried out at Harvard by Brown and his associates of the speech of three small children (referred to as Adam, Eve and Sarah). The first treatment of interrogatives in this work was carried out by Bellugi (1965) and it is this initial description that forms the basis for much of the subsequent work: Bellugi (1967), Klima and Bellugi (1966, 1973), Brown (1968), Brown and Hanlon (1970).

[5] This view stems from a more semantically based notion of underlying structure in which WH-questions were derived from structures like He wants WH-something. Why-questions posed an obvious problem for such a proposal. The current view of the grammar removes such semantic considerations from the description of the operations which permute WH-elements, and rather places them in the domain of the rules which interpret the structures which are the output of such operations.

Children's why-questions in general provide an interesting area in which we can see clearly a potential three-way separation of syntax, semantics and pragmatics. Structurally, these questions behave like other WH-questions in child-language (but, cf. the discussion of Labov in section 2.2.4). Nonetheless, observation of answers continues to reveal that children do not have command of the semantics involved in them. Children do, however, seem to have command of the conversational role of why-questions. It has been suggested that children use these questions to prolong discourse. They seem aware that why-questions rarely elicit one word answers, that each answer they do elicit may be the basis of another why-question, and they exploit such knowledge. Needless to say, this pragmatic knowledge is quite independent of either the syntax or the semantics, and further it seems that the syntax is independent of the development of either the semantics or the pragmatics.

[6] In fact, Wick Miller (1964) suggested that "it might be expected that the child would apply the question inversion to the relative clause. . . ." He, of course, found no examples of such extension.

[7] In a linguistic theory that includes (strict) definitions of rule types and constraints on their application, which are claimed to be part of the hypothesis mechanism, we would have a potential explanation for why such an overgeneralization in the grammar is blocked (cf. section 3.4.2).

[8] Baker (1979) (cf. section 1.4.5), in fact comes to a constraint on deletion from the direction of constraining rules to insure learnability (in the nontechnical sense, recall) given only positive data. This is the no specified deletion constraint:

> Transformational rules may carry out deletion under identity (including the deletion of an original sequence under identity with a transformationally inserted copy), but may not carry out deletions of specified sequences.

(cf. note 20, where Baker points out that this is a special subcase of the strong form of Braine's [1977] "Spelling Prohibition," p. 552).

[9] "Systematic" is included in parentheses because Labov does not explicitly make reference to the notion of systematicity, stressing rather the notion of variation. I assume that he would want to distinguish systematic variation from random variability in that no one would want a linguistic theory to try and provide any account of the latter.

[10] Mayer, Erreich and Valian (1978) countenance the possibility of competing grammars as well.

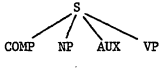
[11] Later in the paper (p. 31), Labov notes that the rule "represents the actual data for the period March-June 1975," a four-month period. However, no other rule is proposed for the remainder of the time, although he claims to have recorded every WH-question uttered during the 17 months.

[12] Strictly speaking, we should think of COMP as SPEC(S); that is, as the specifier of S. Referring to it in such terms, we raise the question of whether COMP itself is a primitive category, or whether there is a more precise characterization of the syntactic category that functions as SPEC(S). We will consider this question in more detail, as well as the related question of what the precise characterization of S should be, in Chapter 3, section 3.1.

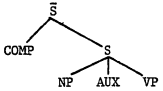
[13] The use of the term trace entails the assumption of the operation of a movement rule; one that I do not wish to imply. More generally, we are referring to the binding of variables or any "gap."

[14] Another fact about COMP that UG apparently leaves to the LAD to deduce is its position relative to the clause. That is, grammars may differ in that they assign to clauses the general shape in either (i) or (ii). In Chapter 3, this possibility and its relevance to the reanalysis of intermediate grammars will be discussed in more detail.

(i)



(ii)



[15] At the stage where this rule is proposed, there is no auxiliary inversion in YES-NO questions; there is a rising intonation contour. The disjunction in (11a), therefore combines an interpretation of Q as an abstract marker, with an interpretation of it as an actually occurring morpheme; what, where, or why.

[16] Klima and Bellugi suggest that a schema such as (i) is appropriate as well: (i) [Nucleus-no].

Of the examples of negative utterances provided as justification for the account that includes (11b) and (i), only one of these shows utterance final negation: (ii) More . . . no. The other nine all exhibit sentence-initial particles. Klima and Bellugi do point out that each sentence in their study represents large numbers of like utterances. Nonetheless, if the utterances are like (ii), the extent to which such utterances reflect grammatical competence is questionable.

However, the issues raised here are wider. For one, this is precisely a place to question the relative systematicity of data gleaned from spontaneous speech. More important, perhaps, is the second issue. Suppose we were to find that utterances such as (ii) were systematic, and did reflect grammatical competence. This finding would not vitiate the argument in the text about the existence of COMP. The role of the evidence from negation in the argument is to demonstrate the tendency of such elements to be placed in sentence-initial position and to pose the question of why this should be so. This role is shared by the evidence from questions--as (11a) shows. The interesting question then becomes not, why do some utterances show negation in sentence-final position, but why do no utterances show either an auxiliary or a WH-element in sentence-final position? The suggestion of an answer to the question may be provided by the relationship of COMP and the positioning of question particles in UG, but in a crucial way, the significant step is the question itself.

[17] Roeper and Matthei (1975) observe further that there is a stage during which children interpret sentences such as (ia) as (ib).

i. (a) The circles are all black.
   (b) All the circles are all black.

They propose that underlying this interpretation is a structure in which the quantifier all is dominated by (the highest) S. In (12) in the text, it seems that the structure revealing interpretation is realized as a surface structure, insofar as we can infer surface structures from utterances. However, no systematic utterances in children's speech of the form All the circles are black, with the interpretation in (ia) have been observed. Further, there must be a distinction in the surface structure placement of the quantifier [as opposed to its placement in the interpretive structure--logical form (LF)] in sentences such as All the circles are black and Only I want candy, which would have the structures in ii(a) and (b), respectively [where iib has the interpretation indicated in (12b)].

ii. (a) $\left[_{S} \left[_{NP}\right.\right.$ All $\left[_{\bar{N}}\right.$ the circles$\left.\right]\right]$ $\left[_{VP}\right.$ are black$\left.\right]$

(b) $\left[_{S}\right.$ only $\left[_{NP}\right.$ $\left[_{\bar{N}}\right.$ I $\left.\right]$ $\left.\right]$ $\left[_{VP}\right.$ want candy $\left.\right]\left.\right]$

Thus, the data in (i) are not clear enough to be relevant to the positing of a sentence-initial category such as COMP, as they stand. Nevertheless, they do make it clear that the development of quantifiers deserves attention with respect to a variety of constructions.

[18] These data are from Bellugi (1969).

[19] As we will consider the problem, it should be stated in terms of establishing the features that define the categories.

[20] We use abductive reasoning in the sense of Pierce (1966). For discussion of the notion of abduction in linguistic change, see Andersen (1973). White (1977) and Chomsky (1980) discuss further how this type of inference applies to language development.

[21] Such a structure entails a rule for the expansion of COMP that is more specific than the rule in (22) in that it must allow 2 branchings of COMP. Note, however, that the branching is not obligatory. COMP may combine single constituents, as in the case of (12) or (13), for example.

[22] It is imaginable that someone might suggest that children have posited a rule moving AUX from COMP to preverbal position. There are two reasons for not seriously entertaining this logical possibility. The first reason is that inferring such a rule would not provide us with any means for accounting for the inability at this stage of WH and AUX to co-occur sentence initially. The second reason is that there is no prototype of such a rule available in UG. To my knowledge, no language has been documented to involve a rule moving tensed elements from sentence-initial position to preverbal position. There is, thus, no reason to suspect that the hypothesis mechanism of the LAD would entertain such a rule. In particular, if the mechanism has already posited a sentence-initial COMP, it would be contradictory to conclude that the mechanism would posit a rule moving AUX away from COMP, given the relationship we outlined between a sentence-initial COMP and leftward movement rules.

[23] But cf. Baltin (1978) for an alternative account of the ill-formedness of such strings.

[24] This is not the case for all such misformulations, however. The intermediate formulation of particle movement, as we shall see in Chapter 4, does involve a structure preserving substitution. However, it may also be the case there that we need not stipulate that the rule is structure preserving, a question we will pursue in more detail in that discussion.

[25] Emonds' discussion of this problem, culminating in the sentence boundary condition (SBC), extends to the problem of complex-NP shift as well.

[26] An interesting proposal, I think, given recent developments in the theory of grammar, is that COMP does dominate a set of features not just exclusively the feature [±WH]. If we can assess the category AUX as a sentence operator, to the extent that it dominates the features ±TENSE/±AGR (cf. Chomsky, Pisa lectures), and if it is the case that AUX does move to COMP in the developed grammar (cf. the discussion in the next chapter) we may consider that AUX shares a feature with WH-elements, quantifiers and complementizers; it has already been proposed that WH-elements behave like quantifiers. Thus, given that all of these elements (or categories dominating them) appear in COMP (but, cf. May [1977] for an alternative analysis of the movement of quantifiers), they may be characterized by some feature capturing the generalization that they are all sentence operators. Then, it may be this feature that characterizes COMP. I have not, however, investigated the consequences of such a proposal, although it is interesting, particularly with respect to the suggestion that COMP is the specifier of S (cf. note 13).

[27] The Subjacency condition as presented in Chomsky (1977, p. 73) states that a cyclic (transformational) rule cannot move a phrase from X to Y or from Y to X in the following configuration, where α and β are cyclic nodes.

$$\ldots X \ldots \underset{\alpha}{[} \ldots \underset{\beta}{[} \ldots Y \ldots \underset{\beta}{]} \ldots \underset{\alpha}{]} \ldots X \ldots$$

The choices that have been considered for cyclic nodes are $\bar{S}$, S, and NP. As (25) reflects, English permits an apparent violation of this constraint on cyclic movement rules. Note that while we are considering only WH movement in the text, NP movement is involved as well:

$$\text{John seems } \underset{\bar{S}}{[\_ \text{ t to be certain } [\_ \text{ t to win}]}]$$

(where t is the trace left by movement of the NP, <u>John</u>).

[28] We are ignoring, for the purposes of this discussion, the Specified Subject Condition (SSC) which is incorporated into the Opacity Condition with the PIC in Chomsky (1980) (note ellipsis in (30)).

The SSC later becomes the Opacity Condition itself, when the PIC is reformulated as the NIC (Nominative Island Condition). This condition, the NIC, in its early formulation, is stated as (i); (where the subject of a tensed clause is assigned nominative case).

(i) A nominative anaphor in S cannot be free in $\bar{S}$ containing S.

Particular conventions on the assignment of case are argued to allow the subsequent simplification of (i) to (ii) (cf. Chomsky 1980, p. 36 ff.).

(ii) A nominative anaphor cannot be free in $\bar{S}$.

29 In fact, if we are to understand completely the information about the development of the grammar that evidence from child language can give us, we need to look at it consistently with more than one suggestion from the theory of grammar. Therefore the suggestion made in the text is not the only one we must keep in mind in an investigation such as the one undertaken here. It may be that the developing grammar of English has, from its earliest stage, a rule, $\bar{S} \longrightarrow$ COMP S, and indeed what must be learned is the correct bounding node for Subjacency, as well as the possibility of lexical exceptions to the generalization in the language. Another possibility is that what must be learned is whether cyclic rules in a given language are subject to Subjacency. One interpretation of Rizzi (1978) suggests this as a valid question to be asked. We consider this question in more detail in Chapter 3. A third possibility, closer to the one here, suggested in Reinhart (1978) is that grammars may vary with respect to their COMP systems; namely, that a language may justify a grammar with two COMP positions.

All of these possibilities exist as competing accounts for the variation across languages with respect to violations of Subjacency, in particular, extractions from WH-islands. Unless we keep all of them in mind, with the kinds of evidence that might suggest which is involved in directing the development of the grammar, we will fail to get the accurate picture we want of how the principles of UG interact in the development of the grammar.

30 The bulk of the work on the internal structure of COMP has concentrated on the ability of WH particles to co-occur with other complementizers. The filter proposed by Chomsky and Lasnik (1977) deals precisely with this question (cf. the discussion on p. 121 of the text). The problem of branching COMP has not, at least to my knowledge, been investigated specifically with respect to WH and other elements; such as AUX or quantifiers, for example. An interesting problem arises, furthermore, with our claim that COMP does branch in the intermediate grammar we are considering. In a branching, COMP (that is, in a structure such as (21)) a negative particle dominated by AUX would not C-command the material in S that is required to be in its scope (cf. the discussion in section 2.3.1). It is interesting

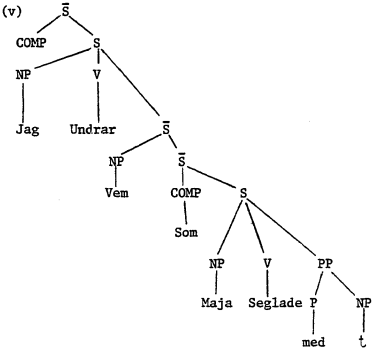
that none of the data in (6), which play a role in the justification of (21) and (24), includes a negative, although the sentences in (i)-(iv) have been attested:

(i) Why not he eat?
(ii) Why not me sleeping?
(iii) Why not...me can't dance?
(iv) Why not me drink it?

Such strings, nonetheless, seem to be very restricted: they are attested with no other WH-form. An interesting possibility is that the failure of a branching COMP to satisfy the C-command requirement triggers the mechanism to provide an alternative grammar in this case, since the subsequent grammar involves a nonbranching COMP. This possibility, of course, assumes C-command is part of the mechanism (i.e., part of UG). There is some evidence for this assumption (cf. Solan 1977). We discuss this in more detail in Chapter 3, section 3.4.2.

A general problem with the assumption that strings like (31) do reflect a branching COMP is that an alternative formulation of WH-fronting proposed in Baltin (1978) does not result in such a structure, but rather in the structure in (v) (cf. Baltin 1978, p. 164).

(v)



That is, WH-fronting is an adjunction to S̄, not a movement to COMP. One piece of evidence which Baltin brings to bear in support of this account comes from right-node raising. A sentence such as (vi) is well-formed:

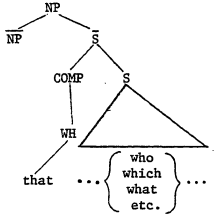
(vi) Jag vet vem, men du vet nog inte vem,
[som har varit har].

I know who, but you know probably not
who, that has been here.

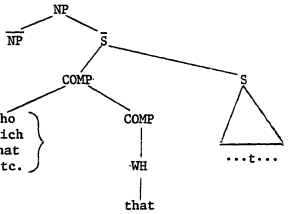
It is of interest, however, that right-node raising evidence in English supports an analysis in which WH is adjoined to S (where COMP is postulated) (cf. Bresnan 1970). So, the alternative account in Swedish does not vitiate the possibility of a branching COMP in principle, although it does raise interesting questions about the correct position of WH-elements.

A question that remains is whether the evidence from Swedish in (31), nonetheless, justifies a claim that the branching COMP is the result of the base rules. In English, the appearance of a branching COMP may result as a consequence of WH-movement in relative clauses, which would map (i) onto (ii).

(i)



(ii)



This result is subject to the filter proposed in Chomsky and Lasnik (1977) for modern English to which we have referred, the Doubly-filled

COMP filter *[WH-phrase complementizer]. Thus it could be that
COMP
the appearance of a branching COMP in Swedish is the result of WH-movement, not the base rules, whatever the correct account of the placement of WH is. For the purposes of the work here, however, it seems sufficient to establish that branching COMP in principle is a possibility allowed by UG. The apparent tendency of the developing grammar to analyze the branching COMP in terms of base rules may be governed by other factors; one being the fact that at the stage we are considering WH, constituents themselves are not placed in COMP by a movement rule.

[31] We do not consider at this point the reanalysis of NEG as a constituent of VP as opposed to AUX, though this reanalysis may depend as well on the analysis of the constituent that is fronted by SAI. No one, to my knowledge, has looked specifically for the possible occurrence--or systematic absence--of errors such as (i) in child language which would reflect the formulation of SAI without reanalysis of uncontracted NEG as a constituent of VP.

(i) Did not Rachael ask Justin to help?

[32] Technically speaking, (32) and (33) allow that SAI can apply to erase WH-elements (34), although independent principles preventing the erasure of lexical material by substitution rules may rule out such an application. These principles, however, do depend on the analysis of the WH-words as either lexical items or grammatical formatives (with only syntactic features; such as DO, for example) in the intermediate grammar.

Chapter 3

"REPLACING" THE INTERMEDIATE GRAMMARS: THE DEVELOPED (ADULT) GRAMMAR, DISSONANCE, AND LEARNING

3.0. Introductory Remarks

In this chapter we have two goals. The first is to propose and motivate analyses for fragments of the adult grammar that are relevant to the learning problem we have set for the hypothesis mechanism by claiming it has constructed the intermediate grammars we proposed in Chapter 2. We will consider the specification of the internal structure of AUX in the adult grammar, the placement of AUX and its role in definition of the syntactic categories S and $\bar{S}$, the placement of COMP in these structures, and finally, the formulation of SAI; in particular, the determination of its output structure. Having provided these descriptions of the "target," and having considered some of their descriptive consequences, a sort of sub-goal, we can turn back to the intermediate grammars and the second goal of this chapter.

That goal is, to set out clearly the two problems that, in a sense, characterize the approach to the study of language acquisition we have undertaken here. The first is the Dissonance problem. This refers to the deviance between the adult grammar and the successive intermediate ones. How can we specify what the differences can be--what goes into the specification? Related to this is the second problem; Learning. What principles make it possible for the hypoth-

esis mechanism to reformulate intermediate grammars in the direction of the next intermediate grammar, and, ultimately, the adult grammar? Further, how do these principles bear on the Dissonance problem? While I have stated the second goal as a specification of questions, it is my hope that the discussion in this section will also provide at least the area in which we might find answers, if not the definitive answers themselves.

3.1. The Description of Relevant Fragments of the Adult Grammar

3.1.1. The Auxiliary--Internal Structure

In recent accounts (Akmajian, Steele and Wasow 1979; LePointe 1980) the auxiliary has been taken to have the shape in (1).

$$(1) \qquad \text{AUX} \rightarrow \begin{Bmatrix} \text{TENSE } \underline{\text{do}} \\ \text{MODAL} \end{Bmatrix}$$

The phrase structure rule in (1) is actually the culmination of several lines of research dealing with the internal structure of the auxiliary, including Emonds (1970, 1976), Jackendoff (1972), Akmajian and Wasow (1974), Culicover (1976).[1] The motivation for the analyses in these works comes essentially from two areas: the syntax of the auxiliary and the morphology of TENSE.

The problems in the syntax of the auxiliary involve the determination of the constituent status of the verbs <u>have</u> and <u>be</u>, originally analyzed as daughters of AUX (cf., Chomsky 1957), as well as their ordering. The reanalysis of <u>have</u> and <u>be</u> was justified both in terms of the theory, and in terms of goals of the actual descriptions. It was observed (Ross 1967) that a number of rules referred to the list

of elements in (2), in particular, subject-auxiliary inversion, and the placement of not in finite clauses.[2]

$$(2) \qquad \text{TENSE} \quad - \quad \left\{ \begin{array}{c} \text{modal} \\ \text{have} \\ \text{be} \end{array} \right\}$$

This recurrence suggested that there was a generalization that current analyses were failing to capture, and that they should, consequently, be re-evaluated, given a theory that requires descriptions to capture maximal generalizations. Second, it was noted that rules, such as VP deletion and VP fronting, for example, variably analyzed particular occurrences of have and be as constituents of AUX and as constituents of VP. The accounts therefore generated them in the VP and proposed rules moving these verbs into the auxiliary. Similarly, NEG particles were argued to be generated as left-most constituents of VP on the basis of their distributional properties with respect to auxiliary verbs (Jackendoff 1972).

Emonds (1970, 1976) further proposed that TENSE reflects the morphology of the AUX system, and proposed (3) as an expansion of AUX.

$$(3) \qquad \text{AUX} \quad \rightarrow \quad \left[\begin{array}{c} \pm \text{ TENSE} \\ \pm \text{ PAST} \end{array} \right]$$

The feature system $\left[\begin{smallmatrix}-\text{TENSE}\\-\text{PAST}\end{smallmatrix}\right]$ characterized the modals can, may, will, shall, $\left[\begin{smallmatrix}-\text{TENSE}\\+\text{PAST}\end{smallmatrix}\right]$ could, would and possibly might. $\left[\begin{smallmatrix}+\text{TENSE}\\\pm\text{PAST}\end{smallmatrix}\right]$ characterizes the affixes -s and -ed; -s being obligatorily deleted in all cases but third person singular. In this account also, DO was oblig-

atorily inserted by transformation in an AUX that carried the feature complex $\left[\begin{smallmatrix}+\text{TENSE}\\ \pm\text{PAST}\end{smallmatrix}\right]$, that is, in an AUX in which there was no modal. A rule of DO deletion operated under appropriate conditions that allowed affix-hopping to attach the TENSE affix to the main verb.

This brief examination of the two lines of work that led to the reanalysis of the category AUX in (1) from the analysis proposed in Chomsky (1957) makes clearer the source of this phrase structure rule which is widely accepted as the standard account of the internal structure of the auxiliary. One modification from its sources that (1) incorporates is the presence of DO in the base.[3] Here again, however, as in the account proposed in Emonds (1976), TENSE and MODAL are in complementary distribution. That is, modals are apparently interpreted as [-TENSE]. This detail, while it captures the morphological facts poses a problem in terms of the role that TENSE plays in the operation of principles of UG. As Chapter 2 argues (cf. 2.4), the status of COMP as an "escape hatch"--allowing unbounded WH-movement--derives from the fact that it is not in the domain of TENSE. Further, the most recent outline of the theory of government binding (Chomsky, Pisa Lectures) assumes that TENSE is a governing category, which, among other things, entails that it assigns (nominative) case to the NP that it governs.[4] In both of these cases, sentences with modals behave as if they include TENSE. Any proposal for the expansion of AUX therefore must take the syntactic role of TENSE into account, as well as its morphological realization. There are two possibilities that suggest themselves; (4) and (5), respectively.[5]

(4) TENSE → { AFFIX DO / MODAL }

(5) INFL → [±TENSE] [±AG]

The difference between (4) and (5) is one of generality. (4) is a rule that is specific to the grammar of English, while (5) is a base rule that could, potentially, belong to UG. [±TENSE] reflects the presence of marking for tense (and aspect as well) in the language while [±AG] reflects the presence of agreement marking for person and number. Because languages clearly differ with respect to the presence or absence of either of these features, as well as to how they are defined, rules of the lexicon would outline the language-specific properties of the category. English, for example, would follow the pattern outlined in (6).

(6)	i.	[+TENSE]	[+AG]	finite verb (and DO)[6]
	ii.	[+TENSE]	[-AG]	modals[7]
	iii.	[-TENSE]	[+AG]	-----
	iv.	[-TENSE]	[-AG]	infinitives and gerunds/ participles

(6iii) is not realized in English, but Chomsky (1980) notes that in Portuguese, the infinitive system does reflect agreement, and (6iii) would, therefore, describe this system. It also seems to be the case that either [TENSE] or [AG] may operate in the language as a governing category. Such a possibility is suggested in the work of

George and Kornfelt (1978) on Turkish in which they argue that [AG] plays a role analogous to [TENSE] for constraints such as the tensed S condition.[8]

What makes (5) particularly interesting as a base rule that is part of UG, with the realization of the features [±TENSE] and [±AG] as what must be learned, is the predictions it suggests for child language. A number of studies suggest that the acquisition of the correct realizations of verb morphology takes some time. As we have discussed (cf. Chapter 2), various analyses have suggested that a copying rule is the first approximation of the traditional rule of affix hopping [Mayer, Erreich and Valian (1978, 1979); Hurford (1975); Fay (1978)]. But, as we noted also, copying may not be the analysis that is correct to infer for the intermediate grammar, because of strings such as (7) (cf. (11), Chapter 2) and (8).

(7) I don't had a nap.

(8) It didn't has any.

The existence of a rule such as (5) with the cross-language variation that exists for the expression of the morphology, and hence the requirement that matrices such as (6) must be learned for each language suggests an interesting account of errors such as those in (7) and (8). Involved in the learning of the matrix in (6) (although not obvious from its outline) is, for one thing, the "organization" of the verbal morphology; are [TENSE] and [AG] discrete morphemes, as they may be in certain languages, or (as is the case in English) are they not discrete, but combined?[9] We would expect to see, therefore, variable realizations of the tense and agreement morphemes, and

interplay with the appearance of DO as well as of modals. The examples of (7) and (8) are suggestive in this direction. We should expect to find, further, strings such as (9) and (10) appearing in the language of children.[10]

(9) John could see "The Empire Strikes Back."

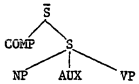
(10) We may ate twenty eclairs.

We will assume, then, for the purposes of our discussion, that the internal structure of the auxiliary is reflected by the rule in (5). In subsequent discussion, we will modify it slightly, but not with respect to the organization of the features. Further, we will assume the existence of a lexical rule spelling out DO to which we referred as well as the rules mentioned in note 5. We will now turn to the problem of the placement of AUX in the developed grammar.

3.1.2. The Placement of AUX in S

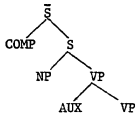
The configuration we will adopt for AUX in the developed grammar is basically that in (11) with some modifications we will introduce in subsequent discussion.

(11)



The other possibility, which we can reject fairly readily, appears as (12).

(12)



The argument that allows us to dispense with (12) originated with Jackendoff (1972), where it was shown that the first auxiliary verb is a daughter of S, but none following the first is. Resting on the placement of sentence adverbs, the evidence supporting the argument indicates that these can appear in all positions available to daughters of S: initial position, sentence-final position, and before or after the first auxiliary verb. Sentences relevant to the placement of the auxiliary appear in (13) (taken from Jackendoff 1977).

(13) a. George will $\left\{\begin{matrix}\text{probably}\\ \text{frankly}\\ \text{certainly}\end{matrix}\right\}$ $\left\{\begin{matrix}\text{have amused}\\ \text{be amused}\end{matrix}\right\}$ the children by the time we get there.

b. *George will $\left\{\begin{matrix}\text{have}\\ \text{be}\end{matrix}\right.$ $\left\{\begin{matrix}\text{probably}\\ \text{frankly}\\ \text{certainly}\end{matrix}\right\}$ $\left.\begin{matrix}\text{amused}\\ \text{amusing}\end{matrix}\right\}$ the children by the time we get there.

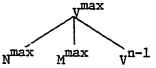
Assuming that the internal structure of AUX is roughly that in (5) in 3.1.1, we see that (11) reflects the placement of AUX most appropriately.

Nonetheless, adopting the outlines of the configuration in (11) still leaves us with a number of possibilities for the relation of AUX to S. We will examine two of these possibilities within the general framework of $\bar{X}$-syntax. The choice between these two possibilities will bear directly on the nature of our proposals for the triggering of the reformulation of the portion of the grammar outlined in Chapter 2.

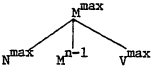
The two possibilities we will consider are in (14) and (15). For the moment we ignore the placement of complementizers in either

configuration. M^{max} in both structures is the label given to the category AUX. Its head, M, dominates the feature bundle specified in (5) above.

(14)



(15)



The major difference between these two structures is the choice of the category that represents S, and hence, is the projection of the head of S.[11] The claim here is that (15) is correct for clausal structure in the developed grammar. An alternative structure, in its essentials like (14), is proposed by Jackendoff (1977). However, there are a number of reasons to question the validity of the claim that V^{max} = S, and that V is the head of S. First a structure such as (14) comes from a phrase structure schema such as (16) which we must assume, given the outline of the framework for $\bar{X}$-syntax presented in Jackendoff (1977).

(16) $V^{max} \rightarrow (N^{max}) - (M^{max}) - V^{n-1}$

The feature that tends to undermine (16) is that it obscures any interdependence between subject and auxiliary. The generalization that the presence of a subject correlates strongly with TENSE in S has long been a part of the literature. In the transformational

framework, it can be traced at least as far back to Kiparsky and Kiparsky (1970) where it is pointed out that "non-finite verbs, particularly infinitives, come about when agreement does not apply. Infinitives arise regularly when the subject of an embedded sentence is removed by a transformation, or else placed in oblique case" (pp. 159-160). Further, in both the framework of "On Binding" (Chomsky 1980a) and of Government Binding (Chomsky 1980b), the assignment of nominative case to NPs in subject position (thus the presence of lexical subjects) is governed by TENSE. That is, the direction of the implication has become clearer: an NP in subject position in an S must be "governed" (a relation that is strictly defined in these frameworks) by TENSE. When the NP is not so governed, it cannot be assigned nominative case. Hence, if there is a nominative subject in S, there must also be TENSE.[12] There is no empirical support for the implication in the other direction: if there is TENSE, there must be subject. We need only look as far as languages such as Spanish or Italian to see that this is the case. These languages optionally dispense with subjects under certain conditions in tensed clauses.

The observation that TENSE assigns Case in both the "On Binding" and Government Binding frameworks suggests further support for the generalization reflected in (15). All other categories assigning case in these frameworks are heads (albeit, lexical heads) of categories; V (head of VP, of course, not S) and P, for example. Assuming Tense is the head of S allows the generalization of this observation.[13]

Work by Judith McAnulty ("Binding Without Case" 1980) independently proposes that TENSE is the head of S, although the structure

she proposes (unpublished notes) differs from (15).[14] A major positive result that she discusses is the possibility of replacing the Tensed-S condition with a generalization of the Left Branch Constraint. The Tensed-S condition, as we discussed in Chapter 2, accounts for the observation that the subject of a tensed clause is opaque. Both (17) and (18), hence, are ill-formed; (17) by the uninterpretability of the trace left by a movement rule,[15] and (18) by the uninterpretability of the base generated reciprocal anaphor <u>each other</u>.

(17) *The men were expected $[_{\bar{S}}$ that $[_S$ Δ would win]].

(18) The men expected $[_{\bar{S}}$ that $[_S$ each other would win]].

The Left Branch condition (Ross 1967), as reformulated in Emonds (1976) is stated in (19) (p. 185).

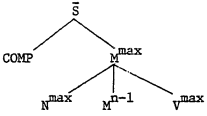
(19) No syntactic element to the left of the head in an NP or an AP can be reordered out of this larger constituent by a major (non-local) transformational operation.

He suggests a generalization of the condition in a footnote (note 10), which is stated here as (20)

(20) No syntactic element on the non-recursive side of a cyclic phrase node H'' [read H^{max})] can be reordered out of this larger constituent by a major transformational operation.

If we assume a structure such as (21) and the cyclicity of $\bar{S}$, ignoring for the moment the exact nature of COMP and $\bar{S}$, and we assume that M^{max} is the head of S, we can see the substance of McAnulty's proposal:

(21)



If the details of such a proposal can be worked out in a way that is consistent with (20), the assumption that Tense is the head of S becomes more attractive insofar as it allows the maximization of what seems to be a more general constraint than the Tensed-S condition.[16]

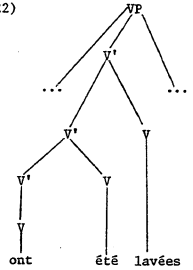
A further advantage of assuming (15) over (14) is that the former, but not the latter, allows us to maintain the strongest hypothesis: that all major categories N, V, A, P, and M share a uniform maximal projection, regardless of what that projection is. Given such support of the proposal in (15), including the arguments McAnulty adduces for the proposal, in its essentials equivalent to (15) (but cf. note 14 above), we will assume that it correctly reflects the position of AUX in the developed grammar.[17]

Before going on to the question of the placement of COMP and the nature of S̄ in the developed grammar, it seems appropriate to discuss briefly some plausible descriptive consequences of assuming a structure such as (15) in which S is the maximal projection of M in developed grammars. We will speculate on the inclusion of a rule of verb-raising in French. The goal is to suggest that this rule is in fact a more general instantiation of the rule of <u>have</u> and <u>be</u> raising in English.

It has generally been argued that French has no class of modals that would be lexically inserted into a base configuration in the

manner assumed for the grammar of English. Thus, the category M^{max} has as its head in the base only the feature bundle INFL. A further proposal has been made for the structure of the VP as it is represented in (22), for a sentence such as (23) (Emonds 1978).

(22)



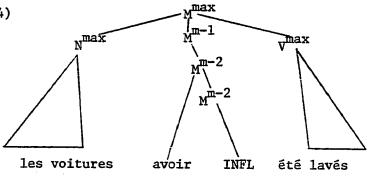
(where "..." indicated optional material)

(23) Les voitures ont été lavées chacque mois.

'The cars were washed every month.'

An interesting alternative is that the derived structure of a sentence such as (23) is (24), prior to the application of morphological rules spelling out tense and agreement.

(24)



By (24) I wish to claim that there is a rule moving the leftmost V into the category dominating INFL.[18] More precisely, the rule adjoins the verb to the left of INFL. This claim extends to all finite verbs. Thus this rule of verb-raising places the verbs in sentences such as (25) into M^{m-2} as well.

(25) Jean détéstera les conséquences de cette proposition.

'John will hate the consequences of this proposal.'

The internal structure of V^{max} may, in other ways, remain equivalent to the structure in (22).

One consequence of suggesting a structure such as (24) is that it provides a natural account of the clearly non-constituent status of the finite verb <u>avoir</u> (ont) and the past participle in (26).

(26) Ils ont sans doute presque tous presque tout trés bien compris.[19]

'They almost all undoubtedly understood almost everything very well.'

Further, it seems that a structure such as (24) allows a simplification of the description of the placement of object clitics, whether by transformation or base rule. It does not need to be stipulated that they form a constituent with the left-most V. Rather they are adjoined to the left of M^{m-2}, and form a constituent with that category. Subject-clitic Inversion too becomes a right adjunction of the subject-clitic to M^{m-2}. Thus we have strings such as (27) from strings such as (28), where we assume <u>me</u> <u>le</u> (l') is part of M^{m-2}.

(27) Me l'as tu donné? (28) Tu me l'as donné?

'Did you give it to me?'

Essentially, the arguments that Kayne (1975) proposes for concluding

that clitics form a constituent with V can be stated equally in support of the claim that the constituent is M^{m-2}. Third, there is a process referred to as Aux-Deletion which generates strings such as (29) (Kayne 1975, p. 95 ff.).

(29) a. Paul m'a bouscoulé et _____ poussé contre Marie.
b. Paul l'a insulté et _____ mis à la porte.

Whether the process is a deletion, or the empty nodes are base generated, it seems a simplification to state the generalization on M^{m-2} rather than to stipulate a specific V (the leftmost) in the process. Finally, assuming the existence of the category M^{m-1} simplifies the rule of L-TOUS (Kayne 1975; Klein 1976; Ruwet 1977), a local rule (in the sense of Emonds [1976]) which moves a quantifier to the left over strings of V. In all discussions of this rule, mention had to be made of the fact that it could not move material over the leftmost V; the quantifier could not occur between subject and the first V. For the most part this fact was attributed to the general principle that the quantifiers could not occur in non-adverb positions, and this position is not open to adverbs. However, that is a generalization that is also stated in the base, where quantifiers are generated in adverb position (cf. Klein 1976). Stating it as a context in the rule of L-TOUS, hence, becomes redundant and complicates the rule. With M^{m-1}, all that must be stipulated in the rule is the generalization that the quantifier moves to the left over V. It will thus never move over M^{m-1}.

Some details of these proposals remain to be worked out, in terms of the levels of M that are crucially involved in the analysis. Hence, the proposals are here as speculation on the consequences of

assuming a structure such as (15). It is important, nonetheless, to point out that the proposals here are not direct consequences of assuming that M^{max} is equivalent to S, and thus that M is the head of S. Rather they interact with that assumption in an interesting way: they propose a plausible explanation for why a number of scholars may have reached the conclusion (which, it is claimed here of course, is incorrect) that $S = V^{max}$ and V is the head of S. It is necessary to work out the analyses sketched above in French involving the proposed generalization of the rule of "verb-raising," which seem to have independent motivation in principle on the grounds of moving in the direction of broader generalization and greater simplicity. If we work them out in a structure such as (14) as opposed to (15), we are left with an interesting question. Is it possible that V was assumed to be the head of S not because of any head-like properties associated with the category itself, but because, in fact, verbal elements may be moved to the category (M) which does have independent support as the heads of S. That is, somehow, certain properties of V (appearing in the position of M) obscured the fact that TENSE is the head of S.

For clarification, let me state this point in a slightly different way. The original motivation for proposing V^{max} as equivalent to S was to allow the generalization of certain grammatical phenomena across NP and S, such as the notions subject of and object of, originally discussed in Chomsky (1970), for example. Thus, the motivation for proposing V^{max} as equivalent to S was not the observation that V had distinct head-like properties with respect to S. In fact, when attention is focused on the question of whether V has such head-like

properties, the answer is ambiguous. Two of the most obvious properties of lexical heads are obligatoriness and the ability to subcategorize complements. While V appears to be obligatory, it is not clear how to interpret this obligatoriness--a question we will consider in a moment. Further, V does not subcategorize any constituents that are immediately dominated by S. While this observation by itself does not constitute a counterargument to the proposal that $S = V^{max}$, it points out a distinct lack of independent support for its presumption.

Returning to the question of the obligatoriness of V, we can consider one possibility and its consequences. First, the apparent obligatoriness of V in a sentence does not by itself justify assuming that it is the lexical head of S, and hence, that $S = V^{max}$. The very assumption that V is obligatory in S can be due to the interference of the actual head of S. That is, the morphological facts--that TENSE (and/or person-number agreement, more generally) is generally attached to V, and that these features are obligatory--may well have lent support to the assumption that V is itself obligatory with respect to S, and that there is support for concluding that it is the lexical head of S. Further, analyses including rules that actually move the V into the M position such as <u>have</u> and <u>be</u> raising, and generally, as in French, verb-raising, as part of this morpho-syntactic process, suggests an additional source for the confusion of V with the actual head of S, or at least an explanation for the assumption that V is obligatory with respect to S, which underlies, in part, its identification as the lexical head of S.

The confusion of the category moved, (V), with the category to which it moves, (M), as the head of S may be easier to understand in English where there is a very restricted form of the verb-raising rule: it only applies to have and be, and there is a rule of affix hopping that can adjoin material from M^n to V. French has no motivation for this rule of affix hopping (TENSE-movement) (Emonds 1978, citing Jean-Yves Morin); that is, it has no rule moving the (non-lexical) head, INFL, out of M^n into V^n. As well, the generalization of the rule of V-raising in French, if correct, seems to lessen the confusion of the two categories.

These speculations about the interaction of the proposals for a rule of verb-raising and the confusion of V with M as the head of S lend support to the proposal of the latter insofar as it suggests a source for this confusion. We will continue to assume (15) for the developed grammar and will see, in 3.2, how the hypothesis mechanism, given the generalization that $S = M^{max}$--i.e., TENSE is the head of S--interacts with other principles to trigger the "unlearning" of the grammar outlined in Chapter 2. We will turn first to the placement of COMP and the nature of $\bar{S}$ in the developed grammar as these will be argued to interact with (15) in the triggering process.

To summarize briefly, the essential character of certain of the arguments presented here to support S as the maximal projection of M is that it is incorrect to claim that S is the maximal projection of V. Further, I have tried to suggest that the latter claim is, in part, the consequence of the misinterpretation of phenomena relating to the status of M as the head of S, and that certain of these

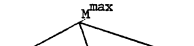
phenomena are related to the rules of verb-raising (in French) and have and be raising in English.

Implicit in the assumption that $S = M^{max}$ is the claim that S is an exception to the general X-bar framework, insofar as it has a non-lexical head. In all other cases, major syntactic categories are the projections of the lexical categories N, V, A, and P.[20] However, there have been acknowledgments of the exceptional nature of S with respect to the X-bar framework (e.g., Emonds 1976 and Hornstein 1977) to the extent that it is not the projection of some lexical category, in particular, V.[21] Thus, it is plausible that S is part of the X-bar framework inasmuch as it does exhibit a symmetrical projection, but is an exception in that the head of S is non-lexical.[22] We will assume, therefore, that S is the maximal projection of M, and that the head of S is not lexical, but is a grammatical formative with the feature composition outlined in (5) and (6).[23]

3.1.3. $\bar{S}$ and COMP

The next problem facing us is to determine the position of COMP in relation to a structure such as (30).

(30)

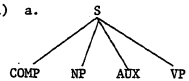


This is a crucial problem for us in two respects. For one, it bears directly on acquisition problem dealing with the interaction of WH-fronting and SAI in each of the ways that the problem can be stated, as we shall see in Section 3.2. Second, the problem is a

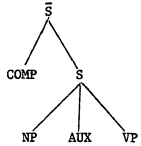
general one in the literature. As it is stated there, the problem has basically two parts. The first is simply where COMP is situated in the phrase-structure tree. The second deals with the internal structure of COMP, a question which was raised in Chapter 2 (cf. note 30), in part, but more completely involves issues such as whether COMP is a prime in the theory or is a cover symbol for a set of nodes that play a more general role; or whether in fact, there is any category (or set of categories) dominated by COMP at all, and COMP is just an artifact of the way constituents move, given the requirements of the binding of the traces of the moved elements.

We have already raised the problem of determining the placement of COMP and its internal structure with respect to the developing grammar (cf. Chapter 2, pp. 14 ff., and note 15). In that discussion, the question of the placement of COMP with respect to S is posed, and it is suggested that UG allows two structures for the placement of COMP in S, shown in (31).

(31) a.



b.

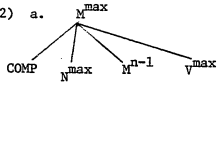


It is the goal here to pursue the consequences of making such a claim, with respect to both its descriptive implications and its implications for the hypothesis mechanism, given it characterizes the "initial

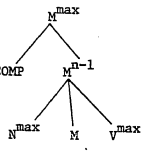
state"; that is, that (31a) and (31b) are available to the hypothesis mechanism.

First, however, the claim must be made within the system where S is a projection of M, and the consequences must be examined from within that framework. Two possibilities that suggest themselves are in (32).

(32) a.



b.



(32a) is the analog of (31a), and reflects the assumption that $S = M^{max}$. In (32b) the assumption is that $\bar{S} = M^{max}$.

Apart from the original motivation for (31b), proposed in Bresnan (1970), that S as well as COMP-S must be analyzed as a constituent for phenomena such as right-node raising, the reason for maintaining the structure in the literature has come from the requirement that COMP be external to the c-command domain of M (Chomsky 1973, 1976). We can interpret recent developments in the theory as requiring, by extension, that COMP be outside the governance domain of M.[24]

Another potential reason for assuming that COMP must be outside the governance domain of TENSE comes from subcategorization facts that hold (in English) between matrix and embedded clauses. This set of factors also provided support for the introduction of COMP into the base (Bresnan 1972). In English, verbs may subcategorize for the com-

plementizers in their embedded clauses. Independent of this fact, a general restriction on subcategorization, proposed by Emonds (1979) holds, appearing here in (33).[25]

(33) Subcategorization Restriction

> If X (= N,V,A,P) is the lexical head of a phrase in a subcategorized complement to Y, then Y cannot (non-idiomatically) subcategorize material in X other than X.

Given this restriction, if $\bar{S} = M^{max}$, then COMP could not be subcategorized for by matrix verbs. Thus, if the restriction holds, and the generalization about the subcategorization relationship between matrix verbs and COMP is correct, COMP cannot be dominated by a projection of M.[26] Before we question whether (32) is part of the initial state; more properly, that UG, which we claim characterizes the initial state, allows either of (32a) or (32b) to be entertained as hypotheses in the developing grammar, we need to establish whether in fact (32b) is adequate for the grammar of English, given the behavior of certain phenomena, in particular, WH-fronting, with respect to the relevant principles of UG. That is, is COMP outside the governance domain of M?

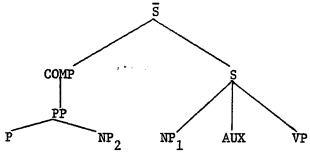
In principle, the answer seems yes. $\bar{S}$ is a projection of S in both accounts, yet COMP in (31b) is assumed to be outside of the domain of AUX in all accounts that we have cited. There is one analysis that does question, if implicitly, whether COMP is outside the domain of S in such a structure. Reinhart (1976) proposes an extension of the definition of c-command, seen in (34) in order to account for the interpretation of the pronominal anaphors in the sentence in

(35). The structure she gives for the relevant elements of these sentences appears here as (36).

(34) Node A c(onstituent)-commands node B if the first branching node α_1 dominating A either dominates B, or is immediately dominated by a node α which dominates B, and α_2 is of the same category type as α_1.

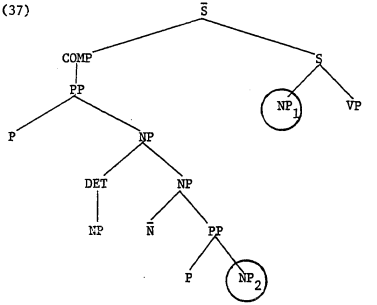
(35) a. *In Ben's picture of Rosa, she found a scratch.
b. In Ben's picture of her, Rosa found a scratch.

(36)



Given the structure in (36), and the refinement of c-command, the NP Rosa in (35a) cannot be interpreted as the antecedent of the pronominal NP her, because NP_1 c-commands NP_2. Of course, drawing this conclusion entails concluding that the remaining daughters of S, including AUX, also c-command COMP. As we have seen, this latter conclusion would vitiate one of the basic motivations for the structure in (31b). In fact, allowing for the moment that these PPs are in COMP[27] we can see that the facts of (35) may follow from the rather exceptional nature of these "picture NPs," which have a structure somewhat more complex than (36) suggests, as we see in (37), which represents approximately what the details of the structure of the NPs in (35) should be.

(37)



Compare (35) with (38), for example, involving PPs with a structure closer to that in (36),[28] and in which there is no contrast in the interpretation of the pronominal anaphor when it appears in the position of NP_1 or NP_2.

(38) a. In his latest exploit(,) Ambrose spends some time at Harvard.

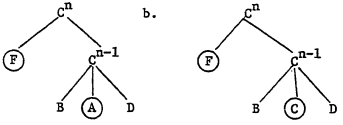
b. In Ambrose's latest exploit(,) he spends some time at Harvard.

Thus, the facts in (35) do not require an extension of c-command and the subsequent abandonment of the assumption that COMP is outside of the c-command (and governance) domain of the daughters of S.

However, although this particular solution to an analytical problem carries over to the framework with (32b), it is still not so clear that we can conclude that COMP in such a structure is outside the c-command or governance domain of M. This is because in (32b), while not in (31b) we are more clearly dealing with a COMP that is dominated

by a projection of the head. In a structure such as (31b) we aren't faced with this problem. The question becomes, then, one of determining whether there is a difference in the way we interpret structural relationships such as c-command or governance in structures like those in (38), where the crucial elements are circled. In fact, it is possible that we should ask whether structures such as (38a) should be allowed at all (since C^{n-1} has no head), and, hence, if (31b) begs the question of determining whether COMP is outside the domain of TENSE.[29]

(38) a. b.

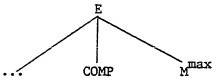


If, for these reasons, we do assume that M^{max} cannot be identical to $\bar{S}$, it is then necessary for us to determine the category dominating COMP and M^{max}, assuming that M^{max} = S. A proposal has been made in the literature (Banfield 1973) that root clauses are dominated by the node E, which dominates the category <u>Expression</u>. E is, with the exception of the conjunction of E (and most likely where root clauses can be embedded [Emonds 1976]), claimed to be non-recursive. In (39) are examples of strings which are expansions of the category.

(39) a. <u>Ah yes</u>, I remember it well.
b. <u>Judas Priest man</u>, we can do better than that.
c. <u>Damn the IRS</u>.

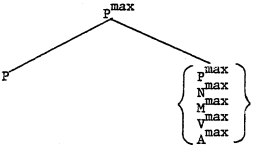
The structure for non-embedded clauses, then, becomes (40), where "..." would contain the underlined material in (39).

(40)



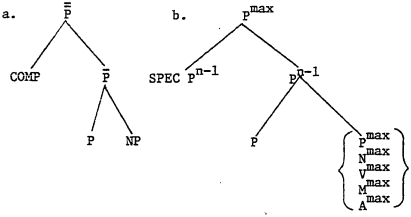
Given that E is peculiar to root clauses in this account, the need arises to determine what might be a likely proposal for the structure of embedded clauses. Such a proposal does exist. On independent grounds, Emonds (class lectures 1979) has proposed that the source of clausal recursion is P^{max}. In the account for which he argues, the category P^{max} includes the feature [±WH]. A generally made observation is that a number of complementizers coincide with the lexical category P, hence the possibility that embedded clauses would occur as M^{max} in a structure such as (41).

(41)



Alternatively, in the description of P^{max} and its possible status as a binding node, van Reimsdijk (1978) proposed that the category may be expanded as (42a), which may be generalized to (42b), in which SPEC P^{n-1} might include COMP as well as the underlined elements in (43).

(42) a. b.



(43) a. John drove three miles up the freeway.
b. Her gaze fixed on the creature that crawled out from under the rock.

Suggesting P^{max} as the source of clausal recursion while maintaining (40) as the source of root clauses, has a number of consequences. For one, it may well be the case that in fact the only elements that COMP can dominate are those having the feature [+WH]. Since P, or SPEC P^{n-1} can dominate other constituents as well as COMP we would have an account of the assymetry observed in the range of complementizers that can occur in embedded clauses and the appearance of only [+WH] elements in root clause COMP position.

Pursuit of the consequences or potential objections to such a proposal is not within the scope of the work here. It is presented only as a suggestion for an account of the assymetry between root and embedded clauses, given the former are dominated by the essentially non-embeddable node E. However, what is important in this discussion is the requirement that COMP be outside the domain of M in the grammar generating structural descriptions of adult English in a theory of grammar outlined in Chomsky (1973, 1977). That is, in a theory that

includes the Tensed-S or Propositional Island Constraint (PIC).[30]

3.1.4. SAI

Although the correct formulation of Subject Auxiliary Inversion is a crucial problem for us here, it is interesting that the rule has traditionally not received a great deal of direct attention in the literature. In most treatments, this rule, generally held to play a role in the structural description of sentences such as those in (44) (i.e., YES-NO and WH questions), has been assumed to be formulated as in (45).[31]

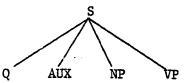
(44) a. Was Marlowe the only operative in San Francisco?
b. Will Ronald sustain this perverse popularity?
c. Have you any wool?
d. Do all things come to those who wait?
e. Who left this mess?
f. Who did Elizabeth abandon?
g. How tall are you?
h. When does the Left Anterior Descending artery become dysfunctional?
i. Why don't babies sleep more?

(45) (Q) NP AUX X ⟹

1 3 2 Ø 4

The rule was, further, generally assumed to have sister adjoined the moved AUX to NP, giving rise to the structure in (46).

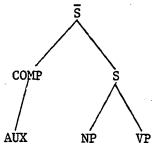
(46)



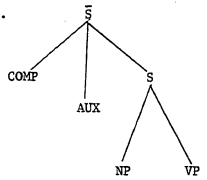
With the introduction of COMP, the problem of the placement of the fronted AUX became somewhat more complicated. In addition to the

possibility in (46), which still remains, even with the reanalysis of Q as COMP, both of the possibilities in (47) exist.[32]

(47) a.



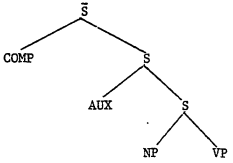
b.



It is likely that a constraint on the available landing sites for movement rules, such as the one Baltin (1978) proposes, rules out (47b) as a potential output. He proposes that moved constituents may be adjoined only to the left-most or right-most boundaries of the categories to which they are joined.

The simplification of the formulation of transformational rules in the direction of "move α" does nothing to mitigate the indeterminancy of the output structure. A rule such as "prepose AUX" might move it to any of the positions in (46) and (47) as well as to the position in (41), given Chomsky adjunction.

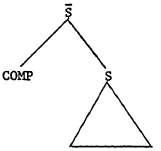
(48)



We will, toward specifying the correct derived structure in the developed grammar, first consider a number of arguments in the literature which favor (47a) as the output of SAI. The first and best known comes from Williams (1974).[33] He notes that the sentences like those in (49) are ambiguous, the two readings relating to the scope of negation. The wide scope interpretation is generally given the structure (50a), while narrow scope has the structure in (50b), where the <u>because</u> clause is not in the domain of the negation; i.e., the reading with comma intonation.

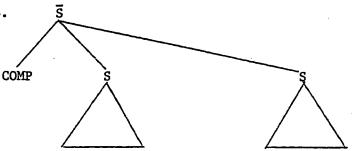
(49) John didn't kill his wife(,) because he loved her.

(50) a.



John didn't kill his wife because he loved her.

b.



John didn't kill his wife

because he loved her

The crucial observation was that the ambiguity disappears in the question, shown here in (51).

(51) Didn't John kill his wife because he loved her?

Williams (1974) suggested that a formulation of SAI as a movement into COMP would explain the loss of ambiguity, because with AUX in COMP in either (50a) or (50b), the because clause would remain in the domain of the negative following SAI, and hence only the wide-scope reading would be available.

Another account that is less well known does entail that AUX moves to COMP not strictly in terms of the topographical position of COMP, in $\bar{S}$, but rather in terms of the requirement that in YES/NO questions, AUX must be a member of the category COMP. It comes from Fiengo (1974). He proposes a number of conditions on the surface interpretation of phonologically null sequences in gapped constructions. Strings composed exclusively of one or more such elements are designated I.

(52) $\ldots X \ldots [_{\alpha} (\acute{L}) \; I \; (\acute{R})]$ X controls I

In this schema, $\acute{L}$ and $\acute{R}$ stand for elements which contain stress to the left and right, respectively, of the empty string. A number of conditions are appended to the rule. Of these, only one is directly relevant to the argument that AUX is in COMP in SAI structures. It is stated in (53).

(53) If I and R contain M_1 of $[+m]$ of α, $L = M_o$ of $[+m]$ of α

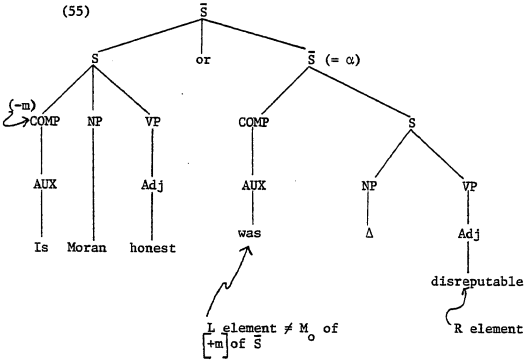
M denotes any syntactic node, and $[+m]$ denotes the daughter of M which is its head, while $[-m]$ is a non-head daughter of M. The symbol M_1 is interpreted as "one or more syntactic nodes," and, correspondingly,

M_o is no nodes. Fiengo defines S as [+m] of $\bar{S}$, COMP as [-m] of $\bar{S}$, and S as having no head.[34] It is the condition in (53) that accounts for the contrast in (54a) and (54b).

(54) a. *Is Moran honest, or was _____ disreputable?
b. Is Moran honest, or was Moran disreputable?

As we can see from the tree structure in (55), (54a) is ungrammatical because L ≠ M_o of [+m] of $\bar{S}$. That is, the left element, COMP, is not empty. It contains <u>was</u>.

(55)



This condition is not proposed exclusively for SAI sentences. It is also claimed to be relevant to distinguishing between pairs such as those in (56).

(56) a. *Today John called from Michigan and yesterday _____ drove from Indiana.
b. John arrived at dawn and _____ quickly went to sleep.

A third account in which SAI is formulated as a rule adjoining AUX to COMP is found in denBesten (1977). One of his stated goals in the paper is to make the heretofore extrinsic ordering of WH-fronting and SAI in interrogatives a consequence of the formulation of the rules, and more generally, of the theory. In the course of reaching this goal, he also proposes to refine the notion "root transformation" (Emonds 1970, 1976), and to extend the definition of the domain of the application of a rule (Williams 1974).

To begin, he argues that root transformations must be divided into two groups; those that are responsible for parenthetical structure and those that belong to the set of COMP attraction rules. It is the second group which concerns us, since it includes SAI, along with NEG constituent preposing, ADV preposing, participle preposing, and PP preposing.

In the course of his attempt to establish the cyclic (in a non-technical sense) domain of movement rules, denBesten cites Williams' (1974) observation, that the domain of SAI as an $\bar{S}$ rule may be established on the grounds that COMP is required in the statement of the rule's affective environment, and that COMP is a daughter of $\bar{S}$, but that the theory does not support the assumption that AUX is placed into COMP.[35] denBesten tries to make this output a consequence of the theory. Roughly, the line of the argument runs as follows. denBesten redefines the domain of the application of a rule so that it includes exactly constants which are affected by a rule, not just those which may effect the application of a rule. Thus if it is the case that SAI is an $\bar{S}$ rule, which, denBesten maintains independently,

given its status as a root transformation, COMP must be affected somehow by the rule, not just present as some sort of trigger. If SAI is assumed to move AUX to COMP, this requirement is satisfied. So while in a sense, the formulation of SAI as a movement into COMP is still an assumption here, denBesten adduces support for the assumption from the theory.

One other analysis which assumes that SAI is a movement to COMP is Hendrick (1979). Here he proposes that reduced questions of the sort in (57) are derived via the rule of free deletion in COMP (FDC, Chomsky and Lasnik (1977).

(57) a. You seen Jane today?
b. Who you seen today?
c. You seeing Jane these days?
d. You happy now?
e. What you doing tonight?[36]

This rule is subject to certain constraints on the recoverability of deletion, which Hendrick outlines in some detail, but which do not concern us here.

There are, then, four independent accounts in which the assumption of an analysis of SAI placing AUX into COMP is crucial, and the conclusion that SAI is so analyzed seems well motivated. However, there is one account of a set of data that poses a problem for such a formulation. Rochemont (1978) notes sentences such as (58).

(58) Who didn't John kill(,) because he was jealous?

Such sentences, unlike the YES-NO question in (51), are ambiguous. Accepting the two available structures for the <u>because</u> clause shown in (50), we need to question the analysis of SAI as an adjunction to

COMP given the structure it makes available for interpretation; the narrow scope interpretation and an analysis which places AUX into COMP seem irreconcilable. That is, there would be no way of explaining the ambiguity of (58), since only the wide scope interpretation would be available.

If SAI is stated as "prepose AUX," might there be some principle defining the way movement rules operate that would ensure precisely the two output structures in (47a) and (48)? That is, is there a mechanism that would result in SAI moving AUX into COMP when COMP was empty, and (Chomsky) adjoining it to S when COMP was filled? Something like Baltin's (1976) principle of landing sites, which we referred to above in connection with (47b) seems relevant here as well.

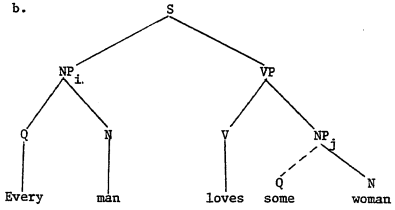
This principle would allow "prepose AUX" exactly two landing sites--those in (47a) and (48), providing we assume that COMP represents the leftmost boundary of S. The Doubly-Filled COMP Filter (Chomsky and Lasnik 1977) would independently rule out any derivations in which AUX were adjoined to a filled COMP.

An alternative to this account suggests itself as well, however, and it comes to mind because of a potential conflict between the foregoing proposal and a constraint on movement rules proposed by Schwartz (1972).[37] The constraint is the Fixed Nucleus Constraint, which states that the head of a phrase may not be moved within its phrase. A rule adjoining AUX to M^{max} (the projection equivalent to S) would appear to violate this constraint. Stating SAI as a right adjunction to COMP would remove the violation.

The almost immediate objections to this alternative are first that it will fail to provide a structural account of the ambiguity in (58), and second, that it will result in multiple violations of the Doubly-Filled COMP Filter. To begin, with respect to (58), we could argue that the syntax need not directly reflect the structure that is input to the interpretive component. May (1977) presents arguments for making such an assumption. For one, he contends that given the c-command relation determines the scope of quantifiers, assuming the surface structure (59) of a sentence such as (59a) (May's 1.15 and 1.29, respectively) is the input to the interpretive component, the ambiguity of such a sentence is left without an account.

(59) a. Every man loves some woman.

b.



May proposes that the surface structures are in fact input to a set of rules, the rules of Logical Form, which, in addition to the rules of syntax, determine the shape of structures that are input to the interpretive component. In this framework we could abandon the assumption that SAI has two distinct outputs given the requirements

of interpretation. We can still maintain that SAI is a rule which adjoins AUX to COMP.

The second objection, that this account would result in multiple violations of the Doubly-Filled COMP Filter, has a response as well. This filter to which we also refer in Chapter 2 (p. 121 ff. and note 30) was initially proposed (along with the rule of Free Deletion in COMP, to which we referred above) to deal with the generation of strings such as (60).

(60) *The man who that I saw is my friend.

All of the unfelicitous strings are of this same general type--they include a WH-word and a complementizer. In fact, the original shape of the filter specified just that combination (cf. Chomsky and Lasnik 1977, p. 434). But this filter in particular does not apply to the COMP of root clauses, since they do not include [-WH] complementizers; these occur only in embedded clauses, as we noted above in our discussion of COMP and M^{max}.

Insofar as these responses to the objections hold, it seems we could accept an account of SAI in which the rule right adjoins AUX to COMP. On the other hand, the one reason for rejecting the first account, in which AUX is placed in COMP when COMP is empty, and Chomsky-adjoined to M^{max} when COMP is filled, was that the operation of the rule in this way would result in a violation of the Fixed Nucleus Constraint. Technically, however, Chomsky-adjunction does not move a constituent around in its phrase, since it creates a new structure. In fact, the second of Schwartz's constraints, the Boundary Attachment Constraint, which states that a phrase moved out of

its phrase cannot attach to anything but the boundary of the next highest phrase, in a sense ensures that any adjunction must be interpreted as Chomsky-adjunction. Such an adjunction would hence comply with the Fixed Nucleus Constraint.[38] The first account, then, remains as a viable alternative, given the foregoing interpretation of the relevant principles of UG. Rather than choosing between these two alternatives here, we will, in the course of the discussion in 3.3, look at what the learning task would be for each of them, since such a question should bear on choices of this type. As well, additional evidence brought to bear in that section will contribute to the determination of the output structure.

3.2. Dissonance and Learning

The foregoing section has been devoted to a presentation of proposals for relevant fragments of the adult grammar. Here we have tried to justify these proposals in light of their relationship to constraints laid out in UG, as well as to outline their descriptive consequences. This discussion is preliminary to a central goal here; that of providing an account, bound by the same principles of UG, of what we may term the "dissonance" between the developed grammar and the corresponding fragments of the developing grammar, the nature of which we discussed in some detail in Chapter 2. The dissonance problem has to do with why it is the case that some particular intermediate grammar is the output of the hypothesis mechanism at a given time, rather than the developed grammar. That is, why might there be mistakes at all in the course of language development? The important corollary to this problem is the question of what are possible as

intermediate grammars; what kinds of mistakes can, and, it appears, must the hypothesis mechanism make in its formulation of a grammar?[39]

The second issue, central to the work here, concerns what we term the "learning" problem. It and the dissonance problem are very closely related, and are often, as we shall see, difficult to separate. The learning problem deals with the set of circumstances that determines the reformulation by the hypothesis mechanism of the developing grammar in the direction of the adult grammar. A major issue involved here is establishing what evidence is required for the hypothesis mechanism to reformulate a misformulated grammar.

At least one potential instance of the learning problem may be an artifact of the fairly nascent status of the work described here. It is possible, for example, that there will be more than one plausible proposal for the description of an intermediate grammar. In such a case we would have to pursue the problem of how the mechanism might reformulate any of the plausible intermediate grammars in the direction of the adult grammar. As our knowledge of the makeup of the hypothesis system is enriched--a process that is as closely tied to the development of the "correct" linguistic theory as it is to the growth of our knowledge about the relationship that holds between language acquisition and the maturation of other, possibly related, cognitive domains--the set of possible analyses of any particular systematic child language phenomena that can reasonably be called intermediate grammars can be reduced, ultimately, to one. Until the indeterminacy is reduced, however, we must consider the learning problem for any proposal relevant to dissonance.

One aspect of the learning problem is, nonetheless, helpful in this regard. Given a fairly clear notion of the kinds of evidence available to the hypothesis mechanism, we can determine whether the mechanism can get from some proposed misformulated grammar to the adult grammar. This is a slightly different way of looking at the learnability of grammars from either the point of view in the work of Wexler and Culicover (1980), or of Baker (1979). In particular, Baker considers the problem from the point of view of determining what are possible analyses in an adult grammar given the requirement that they be learnable on the basis of positive primary evidence alone. Here, on the other hand, we are concerned with assuring that a misformulated grammar can be reformulated in the direction of the adult grammar on the basis of positive primary data alone. In a sense we are imposing a "delearnability" requirement on proposals for intermediate grammars, complementing the learnability requirement imposed in the work such as that cited above.

The "delearnability" requirement should help in limiting what we can propose for intermediate grammars. It is, hence, interpretable as a constraint on the inferences analysts can make. It must, of course, operate in conjunction with the principles of UG, since while both are necessary conditions, neither alone is sufficient to insure that our inferred grammatical descriptions are appropriate. For example, one proposal from the literature which we discussed in Chapter 1 is the S-Initial Hypothesis in Roeper (1978).[40] We have already suggested that claiming that the S-Initial hypothesis is a potential output in the course of acquisition, in fact is a likely output, is

not supported by any principles of UG. There is no formal mechanism to make the acquisition device consider it highly valued as a first approximation because nothing in UG favors such a structure for clausal embedding. On the other hand, the "delearnability" requirement does not necessarily rule out the hypothesis. The evidence necessary for such a hypothesis to be re-evaluated is available in the form of positive primary data. It seems plausible that a child could, from context, realize that a NP in the position of <u>the sheep</u> in sentences such as those in (61) can be the antecedent of the missing NP, and hence re-evaluate a hypothesis that allows only the NP in the position of <u>the dog</u> to be interpreted as the antecedent of the missing NP.[41]

(61) a. The dog pushed the sheep that $[\underline{\quad\quad}_{NP}]$ tickled the horse.
b. The dog pushed the sheep that the horse tickled $[\underline{\quad\quad}_{NP}]$.

In other cases the "delearnability" requirement does restrict the shape of grammars that we may posit as first approximations in language acquisition where the principles of UG do not. For example, it has been argued (Rizzi 1977), that the Subjacency Constraint on movement rules may analyze S and $\bar{S}$ in a language, as is the case in English, or just $\bar{S}$, as is the case in Italian. The distinction is thus "paramaterized" in the sense of Chomsky (1977, 1980), and the correct grammar must be learned in the course of acquisition. The question becomes, then, which of the two hypotheses would the mechanism entertain first for a developing grammar. The answer is not forthcoming from just the principles of UG, since they allow both. The delearnability requirement is what plays a central role here.

If the hypothesis mechanism acquiring English began with a grammar in which Subjacency analyzed only $\bar{S}$, it would require negative evidence to reformulate the hypothesis in the direction of the correct grammar. We should expect, therefore, that the more "restrictive" grammar is the first approximation, given the "delearnability" requirement.[42] A prediction is made for the acquisition of Italian, furthermore. We would expect, if the above holds, to find children acquiring Italian as their native language exhibiting a stage during which they strictly obey the WH-Island Condition. While the Condition holds in English, because of the role played by S in Subjacency, it does not hold in Italian in all cases.[43] In an experimental setting, therefore, we would predict that even children exhibiting competence with the bridge phenomena would either fail to comprehend, or find ungrammatical, sentences like (62) (from Rizzi 1977).

(62) Tuo fratello, a cui mi domando che storie
abbiano raccontato, era molto preoccupato.

'Your brother, to whom I wonder which stories
they told, was very troubled.'

A fruitful experiment along this line would be to test children acquiring Italian using the method employed by Otsu (1980) in his investigation of the instantiation of Subjacency in children acquiring English. The child is told a two sentence story (63), shown a picture depicting the action in the sentence, then asked the question in (64). Given the grammar of Italian, an answer to (64) of a NP corresponding to the English NP <u>a broken leg</u> is not ungrammatical. However, if our proposal for the initial hypothesis about the bounding nodes for Subjacency is correct, there should be a stage at which children will

only respond with the NP in Italian corresponding to a handkerchief, rather than allowing responses with either, which, presumably, an adult grammar of Italian would allow.

(63) a. John is bandaging $[_{NP}$a cat$][_{PP}$with a handkerchief.$]$

b. He is bandaging $[_{NP}$a cat$[_{PP}$with$[_{NP}$a broken leg.$_{NP}]_{PP}]_{NP}]$

(64) What is John bandaging a cat with?

In general, then, we need at least the principles of UG and the "delearnability" requirement in order to limit the range of grammars that we can infer for intermediate stages in the course of language development. Further, stating them makes quite specific empirical predictions which can be tested, as we see in the case of Subjacency.

With both methodological tools in hand, we can turn now to the two problems set forth with respect to the fragments of grammar presented in Chapter 2: the dissonance and learning problems.

3.3. Review of the Developing Grammar

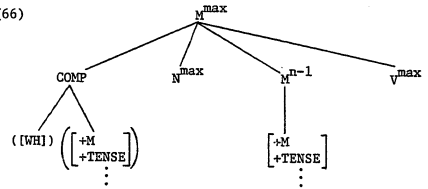
In Chapter 2 we outlined in detail the relevant fragments of the developing grammar and presented arguments, largely based on extensions of principles of UG, which supported the inference of these pieces. Here we will review the order of events that are crucial in the course of the development. In part this review is for ease of reference, but it is also to make it possible for us to see more clearly the interaction of the developing hypotheses and the two problems laid out here: dissonance and learning.

3.3.1. Grammar I (G I)

The earliest stage we consider is the one in which primitive elements of AUX, growing out of NEG elements, appear both sentence initially and in the preverbal AUX position. Other features of this stage, namely sentences such as (65) (which is (6) in Chapter 2) which we argued were not the result of a copying operation, led us to posit the structure in (66) as generable by the base rules.

(65) a. Did you came home?
b. What did you bought?
c. What shall we shall have?

(66)



The structure in (66) differs from (23) in Chapter 2 in that it entails the proposal put forth in 3.1, that S is a projection of the category M, and that the internal structure of the category M consists of a set of features, the precise definition of which must be learned. Hence, there is early confusion as to where certain of these features belong--in the specifier system of M^{max}, that is, COMP, or exclusively as characteristic of the set of grammatical formatives that are M.

The base generated structure in (66) gains additional support from the fact that it provides an account for findings in Gruber (1969). In this work, the following are reported:

(67) a. What do the wheel?
b. What does the truck?
c. Where went the wheel?
d. Where's the a truck?
e. Where's the wheel?
f. Where's the man?
g. What is this?
h. Who's that?
i. Why it go?
j. Where it is?

(68) a. Does it work?
b. Do I make it this way?

(69) a. You fight?
b. See it?
c. You fix it?
d. Me show you?
e. Go round?
f. That go?
g. That broke?
h. That truck?
i. That's a train?
k. It's wagon?

(70) a. Where this guy goes?
b. What this guy do?

In Gruber's discussion, it is observed that the questions in (67)-(69) precede the occurrence of those in (70) which appear eight months later. Gruber himself proposes that the sentences in (67)-(69) do not justify an inversion transformation, although he provides an account quite different from that which would follow from a grammar generating (66).

An account involving movement of the category is not motivated for two reasons. First, maintaining that what is involved is the determination of a set of features, including at least [±AG(reement)]

and [±TENSE], provides a more consistent account for the variable sentence-initial placement of what seems to be the main verb in (67a-h), as opposed to the corresponding placement of an element of M in (68). A possible hypothesis at this point is that whatever element includes the features AG and TENSE can appear as an M.[44] Such a hypothesis is not easily within the province of a movement account, which would require mention of at least two categories, and an analysis of the conditions under which each is moved.

The second reason for rejecting a movement analysis here comes from the subsequent stage, reflected in (70), corresponding to the stage characterized by (1) and (2) in Chapter 2, where there is an inversion transformation in YES-NO questions which is blocked in WH-questions. Positing a movement to account for the sentence-initial verb forms in (67a-c) would make it necessary to explain why the movement was formulated so that it could co-occur with the sentence-initial WH words in these questions, then was reformulated in such a way that prevented its co-occurrence with these same elements in questions such as (70) here and (2) in Chapter 2. In fact, positing a movement to account for the facts in (67)-(69) would require formulating the movement of the finite verbal element so that the movement would be blocked in questions such as (69) and (67i-j), but would operate in questions such as (67a-h) and (68), then explaining a reformulation that almost exactly reverses the pattern.

This second reason for rejecting the movement involves the learning problem as well as questions concerning the empirical data. We need to evaluate the relative availability of evidence that would be

required to reformulate the one movement account to the subsequent one and evidence that would be required to reformulate a base generated account in the direction of an account involving a movement. It turns out that the evidence required for reformulation from a base account to the intermediate transformational account that follows, involving formulation of SAI as a substitution transformation, is more characteristic of the type that we assume to be available to the child.

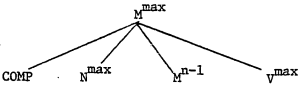
Finally, the sequence of grammars that develops from one which generates (66) to the one reflected in the strings in (1) and (2) in Chapter 2 and in (70) above is itself reflected in the data in (67)-(70). We have first the sentence-initial appearance of inflected verb forms which do co-occur with WH-elements, exemplified by (67) and (68). This is followed by a period, visible in (69), during which there are no occurrences of finite verbal elements sentence-initially--characterized by questions such as those in (69). The final pre-adult stage is the one in (70) and (1) and (2) in Chapter 2. Moreover, the development of each successive intermediate grammar we are proposing from its predecessor, as well as the development of the ultimate adult grammar from the final intermediate grammar has been shown, in Chapter 2, to be subject to the principles of UG and can be shown as well to satisfy our "delearnability" requirement. Showing that this development does satisfy this requirement is the goal of 3.4.

3.3.2. G II and G III

The stage following (66) involves a grammar, G II, that does not generate sentence-initial finite verbal elements, either in the form

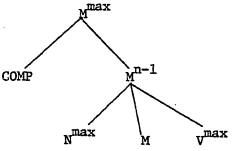
of main verbs (e.g., <u>went</u>) or formatives properly in the category M (e.g., DO). The claim we make is that such a grammar generates structures of the type in (71).

(71)



The structure in (71) is identical to the structure in (32a), above, which, we maintain is available to the hypothesis mechanism in the course of acquisition. The availability of both (32a) and (32b), repeated here as (72), entails the claim that the placement of COMP is parameterized and therefore must be learned. The proposal we are making is that a structure such as (71) is the initial hypothesis tested in the course of acquisition, and that reformulation of the hypothesis is required to go from (71) to (72); the next "step" in the acquisition sequence in our discussion, Grammar III.[45]

(72)



We have maintained throughout this work that the principles of UG favor (71) as the preliminary hypothesis for the input structure to WH

interpretation. It has been claimed, as we discussed, that bounded WH movement is the "unmarked" case, while unbounded movement, such as that occurring in English, is "marked" (Chomsky 1980). Given the framework of UG which includes a constraint such as the PIC, we are claiming that WH interpretation is bounded in a structure such as (71) because COMP, being within the c-command domain of TENSE, cannot act as an "escape hatch."[46]

A further question that we raised is whether the alternative structures for the position of COMP provide plausible description for the distinction between languages with bounded WH interpretation (in the sense of clause-bound) and languages with unbounded WH interpretation. Evidence bearing on the answer to such a question is available in the analysis of child language. A grammar with (71) will have two properties. Not only will WH interpretation be clause-bound in the language described by a grammar generating such a structure, but Subject-Auxiliary Inversion will also be absent. The source of this absence is the constraint on movement rules first proposed in Schwartz (1973), the Fixed Nucleus Constraint, which we discussed in 3.1.4 above. In (71) preposing of M to COMP, the formulation of SAI, would constitute a violation of this constraint because it would be moving a head, M, within its phrase, M^{max}.[47]

The stage for which (71) is proposed as a description, G II, does indeed exhibit a complete absence of inversion, in both WH and YES-NO questions. The two both involve a rising intonation contour, but that is the only marking, aside from the sentence-initial presence of WH-elements in WH-questions. This stage is in fact the one described in

Klima and Bellugi (1973) which immediately precedes the period characterized by the questions in (1) and (2) in Chapter 2, and described here by Grammar III.[48]

The other property, (clause-) boundedness of WH interpretation also appears to be present in the language at the stage described by (71). There are generally no sentences of the type in (73) reported in the literature during this time.

(73) a. Who did Daddy say would read a story?
b. What did Justin think Rachael would like?
c. What time did you say I could stay up until?
d. Where did Mommy forget she had to go?
e. These are the beans that Mommy said I shouldn't put in my nose.
f. There's the policeman that I think will help us across the street.

The presence of such strings in a corpus would allow us to infer that the child had "learned" that WH interpretation is exempt from the PIC, which, we claim, again, is reflected in the structure of (72) (but cf. note 45, p. 216, and the discussion in 3.1.3, above). What then may we infer from the absence of such strings? What we want is to be able to conclude that the grammar at this time (G II) cannot assign structural descriptions to (73). Someone might object that their mere absence from the literature does not allow us this conclusion. This objection does bear some weight. First, to my knowledge no one has looked specifically for the correlation between the absence of sentences such as (73) and the absence of SAI. Thus, their absence from the literature may not allow us to assume their absence from child language. Suppose, however, they are indeed absent from the speech of children during this period. The objection still holds,

since it is possible that there is some constraint on string length that restricts a child's expressive capacity, but has no relation to the grammar. Recall our discussion of the MLU in Chapter 1 in this regard (cf. 1.3.2). It is nonetheless, possible to pin down the question of whether the grammar at this stage does generate structures underlying the sentences in (73) with an experimental methodology similar to that employed in Otsu (1980, 1981), based on the type of experimentation developed in Roeper (1978), Tavakolian (1978), and Goodluck and Solan (1978).

Such an experimental study would involve eliciting children's responses to questions and relative clauses exhibiting unbounded WH-fronting. Using an appropriate set of pictures, combined with sufficient identification of the task, an experimenter could construct stories for which questions and directions such as the following would be relevant:

(74) a. Show me the man who John thinks wanted the ice cream.
b. What did Mary think fell down?
c. Who did Toad say built the dam?

The results of such a study should allow us to discriminate between a child's interpretation of the sentences in (74) as (75), and an interpretation that is consistent with what follows from an adult grammar.

(75) a. Show me the man who likes ice cream.[49]
b. What fell down?
c. Who built the dam?

A child's interpretation of the sentences in (74) as (75) would allow us to infer that the child's grammar did not generate the structures

required for the appropriate interpretation. Further, finding such results during the period when children show no inversion at all, in either WH or YES-NO questions would provide us with the type of evidence we need to argue that the structure in (71) is the right account for this stage. The change that takes place between (71), G II, and (72), G III, should, if our proposal is correct, involve the appearance of phenomena suggesting the operation of SAI, and the correct interpretation of strings such as those in (74). We may or may not begin seeing sentences like those in (73) in the speech of children, but again, their absence is only interesting if it is accompanied by misinterpretation in an experimental study such as the one sketched here. On the other hand, their appearance would give support to our claim about the interaction of the construction of the grammar and the particular principle of UG, and PIC.

It does turn out that during the period where (1) and (2) in Chapter 2 and (70) here, appear, the questions in (76), involving structures of the type in (73) and (74) have also been reported. In fact, they appear in the data reported by Klima and Bellugi, from which (1) and (2) in Chapter 2 are taken.[50]

(76) a. Where the small trailer he should pull?
b. Where the other one Joe should drive?

The sentences in (76), while not precisely like those in (73) or (74), involve phenomena requiring a grammar with the same features as that required for the generation of structures underlying these. Our hypothesis about the succession of grammars, and hence, what must be learned, has support therefore, in the following way. Claiming that the change from G II to G III accounts for the language development

we see in these two periods implies that the acquisition of English does require the learning of the fact that COMP is outside the domain of TENSE, and that it is this property which contributes crucially to the interpretation of a WH-element in COMP--its binding to an empty category elsewhere in the string--allowing the interpretation to be an exception to the PIC. The cluster of properties in the language described respectively by these two grammars: the concurrent absence of SAI and structures underlying (73), followed by the appearance of precisely these two features, supports our drawing this inference.

The final step in the sequence of acquisition we are considering is from G III, in which there is a rule of SAI, but it is formulated as a substitution, to the fully developed adult grammar. The rule of SAI in the adult grammar, as we discussed above in 3.1.4, may be formulated in one of two ways. It seemed there that the principles of grammar involved in establishing the correct formulation of the rule allowed two proposals. The first formulates the rule uniformly as a right adjunction to COMP. The second proposed that the rule results in one of two outputs. In the first case, the rule would be formulated to place M into COMP when COMP is empty--a substitution. In the second, the rule would adjoin M to M^{max-1} by Chomsky-adjunction, thereby ensuring that the Fixed Nucleus Constraint is obeyed. This second output would result if COMP were filled. Hence, it is this second output that would occur in WH-questions.[51] From G III to the adult grammar then, it is the adjunction which must be "learned." Under one proposal for the adult grammar, it would replace the substitution, while under the other it would complement it. We will consider the issues

of this learning in the next section where we discuss the "learning" problem for each successive intermediate grammar from G I through the fully developed grammar.

3.4. From G I to the Fully Developed Grammar: The Learning Problem

3.4.1. G I to G II

The change from G I to G II involves the acquisition of the knowledge that the base rules do not generate any of the features characterizing the category M in COMP, and that the base rules do not generate a branching COMP. The first part of this knowledge may be related to learning which features in fact exhaustively characterize the category M, and are not part of the specifier system of S. One explanation for why this sort of confusion might arise has to do with the "operator"-like status of certain of the features (in particular INFL and TENSE) as we suggested in Chapter 2 (cf. note 26, p. 135). The second part of the acquired knowledge may reflect a parameter allowed by UG. As we also proposed in Chapter 2, it is possible that the branching status of COMP may be parameterized. That is, a language may or may not allow a branching COMP. Modern English has been argued not to, but there are arguments (which we cited above) that Swedish does, and that earlier stages of English did, as Chomsky and Lasnik (1977) discuss in connection with the doubly-filled COMP filter.

With respect to the possible parameterization of the expansion of COMP, the question arises, then, why the first guess of the hypothesis mechanism might be that COMP in English indeed branches, and why we

are justified inferring that this indeed is the first guess. Such questions lead us to look at the delearnability requirement again, as well as the way in which the mechanism might attend to data in the course of acquisition. With WH questions as input, and its hypothesis about the operator status of certain features of M, the mechanism might easily posit a branching COMP, specifically like the one in (66). What sort of evidence, however, would allow the mechanism to abandon such a hypothesis? For one, the determination of the set of features that characterize M, and their role as case assigners instead of S-specifiers could lead to the placement of the features in the preverbal position. As well, we can assume that the input to the mechanism includes data such as the following, where the WH elements and M are not contiguous.

(78) a. I wonder who Daddy loves.
b. Show Mommy what you can draw.
c. Ask Daddy when he is coming home.

Note that not all embedded questions or relative clauses will give this evidence:

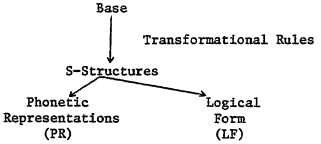
(79) a. I wonder who can eat all her banana.
b. Look at the baby who can crawl.

Strings like (79) would serve to strengthen the initial hypothesis about COMP, whereas those in (78) would lead to a reanalysis of its expansion, with the development of the category M, in terms of its feature makeup

Evidence leading to the reformulation of a branching COMP in English may also come from the interpretation of negatives. A reasonable assumption for the child to make concerning interpretation within

the framework of the theory that has the structure in (80), the framework we are assuming here, is that S-Structure is the input to Logical Form. That is, although the model allows rules of Logical Form (LF) to change the input for interpretation (cf. May 1977, and the discussion above in 3.1.4 relating to this claim), the hypothesis mechanism will assume initially that no other rules are necessary, unless otherwise specified, and that S-Structure and LF are isomorphic. This means that wherever the rule "move α" is required in LF, we are claiming its application must be learned.

(80)



It is generally accepted that the interpretation of the scope of negation involves the notion c-command. That is, a negative element must c-command the constituents in its scope. We have seen this assumption at work in the question of the determination of the placement of M in SAI above and cf. our discussion of COMP in Chapter 2, Section 2.3.1. As the child receives information about the correct interpretation of negatives--all of which can be in the form of positive evidence--he will see that instances of sentence-initial position negatives must c-command material in S. Sentences such as those in (81) would, in context, provide the necessary evidence:

(81) a. Who doesn't want any ice cream?
b. Why won't you finish your asparagus?

A negative element which is the daughter of a branching COMP will not c-command the relevant elements in the sentence. Such a realization could lead to the reformulation of the expansion of COMP. In fact, although it could equally lead to the positing of a rule of LF, I am claiming that a change in the syntax will come first--that is, that the hypothesis mechanism will try to preserve isomorphism between S-Structures and LF. Thus, we might expect, in a language that does arguably allow a branching COMP, evidence for an intermediate grammar that has a COMP that doesn't branch.

We might suppose that these grammars, in which fairly fundamental principles are involved--i.e., the involvement of c-command in the interpretation of negation, the relation of S-Structure to LF--are quite volatile, and appear only briefly. This supposition is borne out, it seems, in English, where evidence suggesting G I is barely noticed in the literature, and we can find it only by piecing together evidence from various sources.

A somewhat simpler account of the early co-occurrence of WH and AUX elements is that the child merely analyzes both as question markers, and, given a preliminary view of English, has noted that question markers are generated sentence-initially. The major objection to such a scenario is that the child does not analyze AUX elements exclusively as question markers, since they do occur preverbally as well at this first stage--in fact they have been preceded by the primitive, exclusively negative reflexes of AUX, i.e., can't and don't, which occur in preverbal position.

A further question that arises is why it might be the case that

(78) is not instrumental in the initial guess about COMP, since such sentences, along with direct WH-questions might in fact lead the mechanism to posit a movement for AUX. Such phenomena have, after all, led linguists to the same conclusion. But the LAD is not the mental equivalent of a theoretical linguist. In the case of the mechanism, the theory, in the shape of principles of UG is, we maintain, part of the biological endowment that is brought to language acquisition. The linguist is not so endowed for his task, which is to determine what the very content of this endowment is. It is reasonable to propose, therefore, that the mechanism plays a significant role in mediating the way data are used in the construction of hypotheses. One way the use of data might vary is in terms of which data are considered most significant. It may be the case that data pertinent to the specification of constraints that are part of UG will play a more important role than data that would influence the positing of a rule. Thus, during the transition from G I to G II, the mechanism cannot posit a rule preposing M because the shape of the grammar will not allow such a rule without a violation of the Fixed Nucleus Constraint. Reanalysis of the grammar, in particular, reanalysis of the placement of COMP, is prerequisite to the formulation of such a rule in compliance with the constraint, as we have discussed. And reanalysis of COMP depends, in the framework outlined here, on the recognition that WH interpretation is not subject to the PIC, in English, and the exemption of WH interpretation from the PIC is due to the positioning of COMP outside the domain of TENSE. What we are saying, then, is that the mechanism cannot posit a movement of

M as it goes from G I to G II, even given evidence such as the alternation of AUX elements in sentence-initial and preverbal position, because the grammar will not allow such a movement. Further, reanalysis of the grammar is not possible until evidence about the unbounded nature of WH interpretation is available. Thus, alternation evidence is "ignored" pending evidence relevant to the shape of the grammar with respect to the PIC, which is a principle of UG.

The natural question to ask is why the hypothesis mechanism doesn't attend to evidence relevant to the PIC earlier. It is plausible that there is a maturational requirement for the condition--not all aspects of the mechanism may be operative from the beginning. Some may require a certain amount of neurological sophistication. As we noted, recent experimental research has been interpreted as supportive of the claim that as soon as constructions relevant to a particular constraint appear in child language, the constraint is operative (Otsu 1980). Prior to the appearance of these constructions at the performance level--if children neither produce nor understand them--there is no evidence that the condition exists. What this suggests is not the nonexistence of the constraint, but rather its inoperative status, as a function of the immaturity of the mechanism. Such a view of language development would be relevant as well to a complete account of Matthei's (1978) results with the Specified Subject Condition, which we discussed in Chapter 1 (cf. Section 1.4.3).

Thus, in the case we are considering here, the PIC, evidence about the unbounded character of WH interpretation in English will be ignored until the mechanism is mature enough to attend to it. The

hypothesis mechanism will, then, reformulate from G i to G II only. It will not re-evaluate the position of COMP, only its branching status. Hence, G II will generate structures in which SAI is blocked, and we will see questions, both WH and YES-NO, with no inversion.

3.4.2. G II to G III

The reformulation of G II to G III does involve the PIC, since, as we have seen, the difference between these two grammars is that the latter includes a rule (SAI), preposing M. Further, we have traced the way in which such a re-evaluation by the hypothesis mechanism of the placement of COMP would occur, involving the PIC. The question we will be concerned with here, then, is why the first approximation for the rule of SAI is a substitution. The question is, in fact, more general, because the claim is that any transformational rule will be formulated as a substitution first by the hypothesis mechanism. Further, there is an implicit assertion that the claim entails. That is that substitutions are somehow "more valued" than adjunctions in the theory of grammar, since we do wish to maintain that the hypothesis mechanism is defined by such a theory.

Why then would a substitution be the guess first entertained by the hypothesis mechanism for a movement rule? A facile proposal would be to suggest that substitutions are somehow simpler than adjunctions. But this is insufficient without some way of making the notion "simple" more precise. In fact there are some ways. To begin with, a rule formulated as a substitution does not create structure as does an adjunction. This is particularly true if we assume, as a number of

scholars do, that adjunction means "Chomsky-adjunction." Such an assumption is made here in the case of the formulation of SAI in the adult grammar. Second, even with a theory of "landing sites" such as that proposed in Baltin (1978), where there is a finite set of possible structural changes in adjunctions, there will be indeterminacy of the derived structure in some cases, as he himself acknowledges in his discussion of WH-fronting and Topicalization.

Following this line of reasoning we can say that in a sense, "simpler" is defined as more restrictive. Formulating a rule as a substitution, the hypothesis mechanism has fewer choices to make, in the case of landing sites, and it has less structure to "worry about" as a result of the output of the rule.

Formulation of a rule as a substitution must, however, also be subject to the delearnability requirement. That is, we must be able to show that in the case where the rule in the adult grammar is not formulated as a substitution, the mechanism will be able to reformulate the rule on the basis of available evidence. There isn't a general type of evidence relevant to the reformulation of any substitution as an adjunction. But, as I attempt to show, for the cases of this type of misformulation I have been able to isolate; SAI here, and the rule of Particle Shift discussed in Chapter 4, evidence does exist for individual cases. It appears that for each case, then, the "triggering" mechanism for reformulation, in the shape of primary linguistic data--which is, in many cases, restricted to positive evidence--must be specified. In a sense, the requirement of such a specification provides us with another test for the correctness of

the intermediate grammar we are inferring: no matter how highly valued a hypothesis may be, if it is not the adult analysis, the mechanism must be able to "unlearn" it, given the relevant data. We will look at the evidence triggering the reformulation of SAI in the next section, as the shift from G III to the adult grammar is involved. Some attention, however, must be paid to the status of WH here first.

We have been assuming that WH elements are generated by the base rules in COMP. The reasons for this assumption are first, that no evidence exists in child language for inferring that children's grammars at the relevant stage generate WH elements in any other position. We have seen that they do not produce either of the echo-type questions in (82).

(82) a. You decided to paint what blue?
b. May you eat how many gallons of chocolate ice cream?

Nor are there any multiple WH questions of the type in (83).

(83) Who polished off which quart of ice cream?

Finally, maintaining that WH elements are generated exclusively in COMP by the base rules gave us an explanation for the absence of copying errors of the type described in Mayer, Erreich and Valian (1978), which we discussed in Chapter 2; there are no copying errors, because there is no movement rule involved.

The question is, then, at what point is there a rule of WH-fronting (assuming that such a rule exists in the adult grammar)? As well, we must ask what sorts of evidence would drive the mechanism to formulate the rule. G III is the appropriate stage of development to

ask these questions, because it is here that WH-interpretation is recognized as unbounded, as we have seen. A plausible question for the mechanism to ask is whether WH-interpretation is subject to Subjacency. We can imagine that while sentences such as (82) and (83) may not be attested in child language data, they are very likely to be included in the data available to the child. They, of course, provide evidence that WH elements can be generated elsewhere as well as in COMP, and such evidence, might lead the mechanism to entertain a movement rule, given G III. The situation here is different than it is in the case of the analysis of AUX and the transition from G I to G II, where we claim that such a scenario is not appropriate to construct. Here, the mechanism is able to interpret data of a different sort. The path is not from (82) and (83) directly to the hypothesization of a movement. Aside from strings such as (82) and (83), input to hypothesis construction, are the fully specified set of constraints, now completely operative, we can assume, given the data we have seen in child language which are relevant to them.

At this point, then, the mechanism has the "knowledge" that WH interpretation is unbounded, and that WH elements can occur in sites other than COMP. As well, we may assume that the mechanism also has knowledge of the Transformational Cycle, given the general content of UG we are maintaining. Together, these pieces of information could lead the mechanism to ask whether WH-interpretation were subject to Subjacency. It is the answer to this question that would determine whether in this case a movement were tested. In these ways, the determination of a movement for WH phenomena is more complex than the

development of SAI. But WH fronting is also, in a sense, "easier to learn" than SAI. Although it requires a more involved set of conditions for the rule to be hypothesized in the first place, the rule is sufficiently constrained so as to make errors almost impossible. In fact no errors having to do either with the interpretation or the placement of WH elements have ever been reported in the literature.

One more question exists. That is, what sort of evidence is sufficient for the mechanism to conclude that WH-interpretation is subject to Subjacency? Positive evidence is available, although it is more of an indirect kind. The mechanism would need to "listen" for violations of Subjacency by the rule. Hearing none, the conclusion that the rule is a movement could be drawn, given the other sorts of evidence we have suggested as well.[52]

3.4.3. G III to the Fully Developed Grammar: The Reformulation of SAI

We have claimed that the evolution of the adult formulation of SAI may involve an addition, rather than a replacement. We have maintained that one alternative for the adult formulation of SAI (prepose M) is as a substitution for COMP when COMP is empty, and an adjunction to S (M^{m-1}), defined by the conditions we outlined. What then does the "learning" problem entail? That is, if we were to assume this adult formulation, how would it evolve from G III?

Given G III, including a rule of SAI, certain co-occurrence data will influence how the rule is formulated. The claim is, that at this stage, the mechanism <u>will</u> be receptive to data in which various elements do co-occur with a preposed constituent of M, where before we

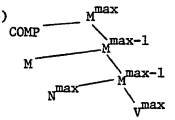
have claimed that it would not. Here, then, the mechanism should attend to WH-questions, as well as to data such as the following, involving other environments for the rule.

(84) a. So endowed with ambition was the princess that she desired more than just marriage to the prince.
b. Seldom do children who don't eat broccoli get fudge sundaes.

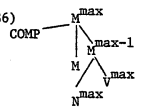
An adjunction, then, would be added to the formulation of the rule.

What would ensure that this adjunction would be formulated with the structure in (85), and not that in (86), which is the analog of (47b).

(85)



(86)



I think it is possible that the interpretation of negatives like the one in (58), which led us to the formulation here may play a role in determining the correct formulation. In our discussion of (58), and of the two possible formulations of SAI (the other being, recall, that the rule is a successive adjunction to COMP), we raised the possibility that the interpretation of (58) may not indeed be crucial to the determination of the placement of M, given the role that movement rules in Logical Form (LF) may play. However, in our discussion of the development of the grammar from G I to G II (Section 3.4.1), we suggested that the initial hypothesis the mechanism makes about the

relationship of S-structures to LF is that the two are isomorphic. Hence, information about interpretation will bear upon the derived structure.

The interpretation of sentences such as (58) as ambiguous is not obvious, and in fact may be argued to be tenuous. In the absence of such evidence we can still rule out (86), given the theory of landing sites of Baltin (1978) to which we have referred, since it stipulates that a constituent must be attached to the left- or rightmost boundary of the category to which it is adjoined. So, once the mechanism is committed to an adjunction, this theory of movement rules would help to reduce the number of possible derived structures. Then, however, we are left with the possibility that the mechanism does reformulate SAI as an adjunction to COMP. While it is correct, I think to maintain that this is not the formulation in the adult grammar (cf. the discussion in note 51), it is possible that the mechanism does acquire this formulation, and may not abandon it until quite late, subsequent perhaps to evidence like the parentheticals. Such indeterminacy, however, is to be expected, I think. It will be limited to cases where the misformulation has either no detectable empirical consequences, or as is the case here, where the consequence (if it is not the interpretation of sentences such as (58)) is very limited in nature, i.e., WH-questions with parentheticals.

3.5. Concluding Remarks

There are two general "morals" we can draw from the detailed path traced here from G I to the adult grammar. The first is that it seems

plausible to assume that data play a variable role at different stages in acquisition. That is, the same evidence may at one time be ignored by the mechanism, as must be the case at the stage of G II, with respect to WH-questions and YES-NO questions, or the data may be "used" in a misformulation, as is apparently the case in G I, with respect to WH-questions.[53] This selective attention to data, we have seen, is dependent on the operation of the hypothesis mechanism in terms of the grammar it is working with at a given time, and its own contents, which may itself develop with the maturing neurological system specific to language.

The second "moral" is that we may end up with an indeterminacy in some cases. The indeterminacy with the exact formulation of SAI may be such a case, although here a case is made for the existence of evidence that would bear on the attainment of the formulation we claim is correct for the adult grammar, albeit limited in availability perhaps. Nonetheless, I believe that in some cases involving peripheral phenomena, individuals in the same linguistic community may end up with different grammars, since the evidence bearing on the choice between two alternative analyses is not available in one case or another, or the evidence itself does not impose a choice. The theory of grammar we are assuming does allow for such an outcome, in fact (cf. note 5, p. 89, Chapter 1). It is left, however, to make precise the notion "peripheral phenomenon" and to set some limits on the indeterminacy.

Finally, it is fruitful to review the goals of this work in light of the discussion in this chapter. The goal is not merely to provide

grammars that will generate the strings we observe in child speech: at one stage questions without SAI, and at another, apparent complementary distribution of SAI and WH-fronting, for example. Nor is it to provide a natural history the development of some construction in the course of acquisition; Questions, for example. Rather, the goal is to suggest how grammars that do generate the structures underlying such strings might be constructed in the course of language development by the hypothesis mechanism, and further, how the mechanism would "unlearn" them, resolving the dissonance between them and the fully developed grammar. In light of this goal, we have attempted an account of the appearance of clusters of properties that would be the consequence of the hypothesization of particular grammars by the mechanism. Thus, we have seen why it would be the case that at a time when there is no SAI, there is also an absence of unbounded WH-interpretation, and why SAI, when it develops, might first be formulated as a substitution by the mechanism. Further, although we assumed that G III evolved to an adult grammar in which SAI involves a substitution and an adjunction, we can see where the structure of the mechanism with respect to the relationship between surface structure and LF--that it assumes isomorphism and must learn where there is a rule in LF--might favor this adult formulation given its effect on the acquisition task.

Notes to Chapter 3

[1] Jackendoff (1972) also cites Klima (1966, unpublished class lectures) as proposing an account of the auxiliary in which have and be were not generated by the base rules as daughters of AUX, but rather as VP constituents.

[2] In early analyses, not-placement was, of course, a transformation. However, the distributional facts remain even when NEG particles are generated by the base rules.

[3] Culicover (1976) proposes an account along similar lines, but in which DO was generated by the base rules as a member of the category modal. The phrase structure rule expanding AUX in that account was, hence, AUX → TENSE M.

[4] This discussion is based on research carried out in configurational languages. Although tense (or INFL, as we shall describe it below) may play a central role in non-configurational language as well, it is possible that the role will not be defined strictly in terms of government, itself a configurational notion.

[5] The accounts of AUX that have been cited, as well as the two proposals here, assume that the grammar also includes rules of DO-deletion and DO-replacement (or raising of HAVE/BE from the VP to AUX). The base rule in (5), further, adapts a proposal for the category INFL in Chomsky (Pisa Lectures).

[6] In a grammar with (5), DO is inserted by lexical rule in the presence of the features [+TENSE] [+AG]. The rules of deletion and replacement outlined in [note 5] remain.

[7] Dialects may differ as well as languages with respect to the realization of those features. It has been noted that Black English does not exhibit (overt) marking for agreement in the finite verb system. Hence (6ii) in this dialect would include finite verbs and modals, while (6i) would include possibly only DO.

[8] This variability suggests that what must be learned as well, is what feature of INFL acts as a governor in a language. To test such a hypothesis, we would need to look at languages in which the two are separable, and then to look for potential errors. The class of errors might include looking for inappropriate case marking for subjects of infinitives in the child language of languages such as Turkish, or to look for evidence of extraction from strings in which

it is blocked in the adult grammar. Evidence such as this might further suggest experimental work that could be done. At this point, I don't believe it is clear, however, how experiments testing comprehension, for example, might shed light on the question.

[9] The clearcut cases of discrete vs. non-discrete would correspond loosely to the traditional classifications of morphological systems as isolating, agglutinating or inflecting.

[10] An interesting gap in the errors that have been reported are those which would involve marking modals as [+TENSE] [+AG] such as (i)-(iii).

$$\begin{array}{ll} \text{(i)} & \text{He coulds} \\ \text{(ii)} & \\ \text{(iii)} & \end{array} \left\{\begin{array}{l} \text{eat} \\ \text{ate} \\ \text{eats} \end{array}\right\} \text{ten Fudgicles}$$

That is, there are, to my knowledge, no recorded errors with agreement marked on the modal. If indeed there is a gap, and no such errors exist we would need to examine what might block such errors. One potential account takes us back to the work of Baker (cf. Chapter 1.4). The acquisition of the auxiliary system does, to a large degree, depend on external evidence--as we have tried to show, given a rule such as (5), and the necessity of learning systems such as (6). If positive evidence plays a central role in such acquisition, we see that the acquisition mechanism has only evidence that modals as a class, take no overt agreement morphology in English, as opposed to the class of verbs, for example. Hence, there is no "room" for it to construct a grammar in which strings with such marking could be generated. The only possible source of errors such as those in (i)-(iii), therefore, would be the mistaken hypothesis that could is a verb. It appears that no child has ever constructed this hypothesis.

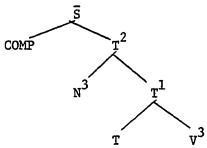
[11] This discussion assumes that S does, in fact, have a head. Work does exist in which the contrary assumption is made, Fienogo (1974), although no argument is made for the assumption, and nothing in that work crucially depends on it.

[12] The term TENSE here is used in loose synonomy with the feature bundle labeled INFL in (5) above.

[13] It may be possible to extend this observation to the head of N^{max}, insofar as it can assign genitive, in the case of possessives. This extension is suggested in work by Judith McAnulty; cf. note 14 for further discussion of her proposal that TENSE is the head of S. A further possibility is that Adjectives may also assign case, at least in some languages (e.g., Russian). Thus, the generalization is that category heads assign case.

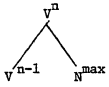
[14] McAnulty (forthcoming) proposes (i) as the structure with Tense as head of S. We ignore the question of whether $C^{max} = C^3$, as well as the relation of $\bar{S}$ to S.

(i)



T^1 is roughly equivalent to Chomsky's "Pred Phrase" (1965). This structure closely resembles the structure in (12), with an alternate label for the highest VP. There is little empirical evidence that would help us decide between (i) and (15), however. One bit of evidence involves the strings (13). Given that there is a rule of ADV placement, its statement on a structure such as (15) is quite general: adjoin ADV to N^{max} or V^{max}. This statement would operate freely to sister-adjoin the ADV to the left or right of these categories, accounting for the sentence-initial, sentence-final, pre-aux and post-first-aux distribution of these elements. Even a more general statement, attach ADV to M^{max}, would, in a theory of movement such as Baltin (1977) proposes, be possible with (15). Neither formulation is possible for (i), however.

One further question to be asked is whether there is a reason to assume that V^{max} (V^3 in (i)) is a complement at the level of T. To my knowledge, none of the lexical items that can be inserted for the feature bundle INFL that would be dominated by T (or M) are strictly subcategorized for by T (or M). In most other cases of this kind, the complements do bear such a relationship; $[\underset{VP}{\ \ V\ \ NP}]$, for example, where we would have the structure



In (15), we can say that V^{max} is an optional complement to M^{max}, which seems to yield the correct results. A problem is that we

probably also want to say that some realization of M, that is, [+Tense ± AGR], strictly subcategorizes for lexically specified N^{max}. This problem could possibly be solved given some refinement of Emonds' (1980) proposal that specifiers are obligatory, and given an interpretation of N^{max} as the specifier of M^{max}, although I don't have a precise idea of how the details might be worked out. This problem extends to (i) as well. McAnulty assumes a rule deleting Tense in the context [PRO TENSE].

15 We assume the trace theory of movement rules as part of the theoretical framework for our discussion. Note that (17) is also ruled out by [the early formulation of] Subjacency.

16 Emonds (1979) ["Word Order in Generative Grammar"] considers the possibility of the Generalized Left Branch Condition.

17 This assumption is also made as part of the framework of the theory in the Pisa Lectures on Government Binding (Chomsky 1980).

18 A similar rule is proposed in Emonds (1978). There the raised V' is a daughter of S.

19 This example comes from Kayne (1975), p. 87, note 18.

20 Strictly speaking, P is not a lexical category, in that it is not an open class. Thus its inclusion already extends X-bar to the specification of projections of grammatical categories.

21 Cf. especially note 23 in Hornstein (1977).

22 It might be claimed that in fact modals constitute a lexical category, and hence, that S is still a projection of a lexical head. Modals do not strictly satisfy the definition of grammatical formative suggested by Emonds. That is, no grammatical rule requires the specification of a particular modal as opposed to the specification of the category. On the other hand, they do constitute a quite small nonproductive lexical class and in this regard are exceptional with respect to the other major lexical categories: M, V, and A. P is, of course, a relatively small nonproductive lexical class, but it is a characteristic of this class that some of its members (e.g., of, to, by, and the like) have been argued to be grammatical formatives inasmuch as they do satisfy Emonds' definition.

Further, there is one argument (Chomsky 1965, Chapter 2, note 9) that modals constitute a lexical class. It comes from the generalization about a rule of conjunction stated as follows: "If XZY and XZ'Y are two strings such that for some category A, Z is an A and Z' is an A, then we may form the string XZand Z'Y where Zand Z' is an A." It is established that A must be a major category, and major categories are themselves defined as projection of lexical categories. In the case of modals, strings such as John can and should manage to change that diaper, suggest that by this criterion, Modals are lexical categories.

The precise character of the class of modals vis à vis its status as a lexical category is a problem that will be set aside here. We will continue to assume that the terminal symbol in M^{max} is the bundle of features crucially including [±TENSE] and [±AGREEMENT]. See also note 20.

[23] In fact, the label M is somewhat arbitrary, and it is possible that INFL (cf. (5) and (6) above, and note 5 and the reference cited there) describes the category exhaustively. However, M does capture the notion of modality that the category does express, even in the absence of overt modals, while the feature matrix captures the distinctions in the realizations of this category that are crucial to the syntax. We will continue to use this label with the acknowledgment that these problems remain.

[24] Where <u>government</u> is defined as follows. "In a structure $[_{\alpha} \ldots \gamma \ldots \beta \ldots \delta \ldots]$ where β is an immediate constituent of α and there is no category properly containing β and properly contained in α, β governs γ providing there is no category properly containing γ and contained in ... . If the same conditions hold between β and δ, similarly β governs δ"(Chomsky 1979).

[25] Emonds cites work by Milner (1976, pp. 41-43) in his statement of (33). Note also that $\bar{X}$ may be generalized to X^{max}. That is, there is no need to assume a particular value for "max."

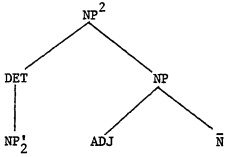
[26] The alternative is that verbs do not subcategorize for complementizers, but they subcategorize for realizations of TENSE, and TENSE in turn subcategorizes for complementizers. In this latter case, there is no need from this point of view for the assumption that COMP is outside the domain of M. Independent evidence that COMP must be outside this domain would be one source of evidence bearing on the choice between these two alternatives.

[27] Regardless of how we interpret c-command and the relation of S and $\bar{S}$, the problem of the placement of these preposed PPs must be determined. They pose an interesting problem with respect to the placement of AUX in SAI in questions such as (i), particularly if an assumption is made that SAI places AUX in COMP (cf. the discussion in 3.3 below).

(i) a. In his latest exploit, does Ambrose Usher spend the whole time at Harvard?
b. In Ben's picture of her, did Rosa find a scratch?

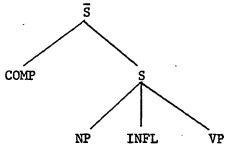
[28] That is, the NP_2 (in (i)) within the larger NP is "higher" (in some sense that needs to be defined) than the circled NP in (37).

(i)



[29] This is a question that is begged, I believe, in recent work assuming INFL(ection) to be the head of S in a structure such as (i), similar to (31a).

(i)



If S and S̄ are primes in the theory, how do they relate to the other major categories, N, V, A, and P? That is, how do they relate to their head, INFL? In what sense are they projections of INFL, for example?

[30] The theory has changed substantially from its definition in the works cited. For one, the PIC has been supplanted by the Nominative Island Condition (NIC), which we discussed in Chapter 2, and both of these conditions have essentially been replaced by the Empty Category Principle (ECP) (Chomsky 1980). This latter constraint on the binding of phonologically null NPs seems to remove the necessity of COMP being outside the domain of TENSE.

The interaction of the data from child language with the theory including the ECP would have different results, in the absolute sense, but it is not the case that such results would vitiate those in this work.

[31] For formulations like (48) see Akmajian and Heny (1976); Culicover (1976); Akmajian, Steele and Wasow (1979). The abstract question formative Q is referred to in Akmajian and Heny (1976).

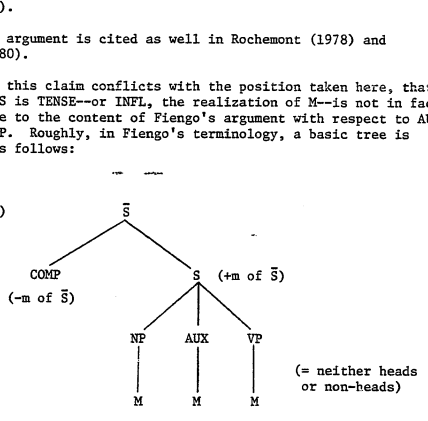
It is, of course, a precursor of COMP; a development traceable from Baker (1970) to Bresnan (1970).

[32] While (47a) and (47b) do not both follow from the assumption that AUX is sister-adjoined to the (subject) NP, they are possible given the formulation in (38), but cf. the text for discussion of Baltin's principle of landing sites, which might rule out structures such as (47b).

[33] This argument is cited as well in Rochemont (1978) and Hendrick (1980).

[34] That this claim conflicts with the position taken here, that the head of S is TENSE--or INFL, the realization of M--is not in fact of importance to the content of Fiengo's argument with respect to AUX being in COMP. Roughly, in Fiengo's terminology, a basic tree is designated as follows:

(i)



[35] Thus, the argument that Williams proposes for assuming AUX is placed into COMP is the descriptive one we here summarized above.

[36] This reduced question also involves phonological reduction; resulting in [ωəčəɑʊɛ̃η . . .].

[37] While this constraint is not strictly part of the theory outlined in Chomsky (1973, 1977) by which we define UG, it certainly does seem to reflect a property of movement rules, and therefore seems reasonable to continue assuming as part of the theory, and, hence, like the principles of UG, part of the genetic endowment we are assuming for acquisition.

[38] The thrust of Baltin's (1978) landing site principles is, in fact, reminiscent of Schwartz's Boundary Attachment Constraint, since the goal of both is to constrain the possible "targets" of movement rules, and they attempt to reach this goal in parallel ways.

[39] When we speak of mistakes here, we are talking in general of misformulated grammars, and not about directly observable errors in the speech of children. As we discussed in Chapter 1, the goal is not to write grammars to describe the speech of children. That, as we noted, turns out to be a fruitless pursuit. Rather, we are proposing, from a cluster of what seem to be systematic observable properties, and the principles of UG, what are possible to infer as descriptions of the child's linguistic competence at the time in question. The task then remains, as the text points out, of determining how the hypothesis mechanism can "unlearn" its grammar, and formulate the adult grammar.

[40] The issue raised in the text here holds equally of Tavakolian's (1977) Conjoined Clause hypothesis, which was also discussed in Chapter 1. The relevance of the S-Initial hypothesis to the "delearnability" requirement was raised in conversation with Nina Hyams, to whom I am indebted.

[41] It is equally important to note that meeting the "delearnability" requirement is not a sufficient criterion for assuming something like the S-Initial hypothesis. One could think of other accounts for the reported interpretations that also meet the "delearnability" requirement but do not require the positing of structures that have no support from the principles of UG.

[42] A similar conclusion has been drawn from work by François Dell (1981) in phonology on the acquisition of optional rules. He suggests the existence of the following as an acquisition strategy: "Whenever there are two competing grammars generating languages of which one is a proper subset of the other, the learning strategy of the child is to select the less inclusive one" (Dell 1981, p. 34).

[43] Rizzi's arguments in Italian are based on extraction of a relative pronoun from an indirect question. He points out that extraction of an interrogative pronoun from an indirect question results in questionable or ungrammatical sentences, but due to reasons having more to do with the interpretation of questions in Italian, and not relevant to the problem of Subjacency.

[44] This hypothesis is probably too loose insofar as imputing it to the child would not account for the general absence of the misanalysis of a wider range of verbs as belonging to the category M in English. J. Emonds suggested that there might be some connection to frequency, in particular, with relevance to "went" in (67c). He points out that in Catalan, go is an aux.

[45] The learning problem for the positioning of COMP must really be stated for either of the two possibilities we outlined in 3.1.3. Here we have just assumed that $\bar{S} = M^{max}$. But there we also suggested that it is possible that this assumption is incorrect, and $S = M^{max}$, while $\bar{S}$ in root sentences is expressed as E. If this latter account is correct, including the proposal for the source of embedded clauses, then the principles of UG would have to make this analysis the consequence of establishing that COMP is outside the (c-command) domain of M in a language. Thus the configurations themselves aren't learned--they're "available." It is the set of circumstances out of which the configurations would fall that must be learned: that the interpretation of WH elements is unbounded, and that SAI is formulated so that it does not violate Schwartz's Fixed Nucleus Constraint. For further discussion of this second circumstance, see the text below.

[46] In the framework we are assuming here, where exemption of the interpretation of WH elements from the PIC depends on the placement of WH in COMP in a structure where COMP is outside the c-command domain of TENSE, the assumption that there is a transformation "move WH" does not follow automatically from this property. The conclusion that there is a rule "move WH" is argued to follow from its adherence to the condition of Subjacency; a constraint distinct from the PIC. A child could "learn" about the exemption of WH phenomena from the PIC before he learned that the rule is a movement (if indeed it is a movement). Hence, even in a grammar such as G III, which we argue below will allow sentences such as (73), all that can be inferred in the framework we are assuming is that WH-elements are generated in a COMP that is outside the domain of TENSE. It is highly plausible that they continue to be generated by the base rules until it is learned that WH-phenomena are subject to subjacency--for example, in English, until the hypothesis mechanism has "contact" with WH-phenomena. Such a scenario of course entails an acquisition framework that assumes a theory of UG in which Subjacency is a constraint on movement; a condition on the operation of a rule. For the view that it is a condition on representations, see Friedin and Lasnik (1981).

[47] The implication here is that there is a correlation between the absence of unbounded WH-fronting and the absence of a rule such as SAI--finite verb fronting in the construction of questions. The implication seems to hold. In Russian, which has been claimed not to allow unbounded WH-fronting, there is no rule of finite verb fronting (≈ M preposing) that might be considered an analog of the rule in English. YES-NO questions are signalled by intonation:

(i) On pri·echal? 'Did he arrive?'
'He perf·come·past'

There is a construction with the complementizer <u>li</u> in which the verb may be fronted, but this is said to be a focus construction:

(ii) priechal li on? 'Did he arrive?'

As well, finite V is not the only constituent that can be preposed:

(iii) mnogo li on pil? 'Did he drink alot?'
much comp he drink-past

[48] Klima and Bellugi describe (71), Grammar II, as the first stage, but, as we have seen, other data, albeit scattered throughout the literature, indicate that this period is preceded by one best described by Grammar I.

[49] In the case of relatives such as (74a), we would also need to determine that children can discriminate between sentences such as (74a) and one such as (i).

(i) Show me the man who thinks John likes ice cream.

Failure to discriminate would indicate that (i) is a possible simplification of (74a), and would also allow us to infer that the child's grammar was not generating the structures which we claim underlie sentences involving unbounded WH-fronting.

We should acknowledge that any experiment like the one sketched here assumes that the constraint crucially involved in limiting interpretation of the empty categories in question, which we are claiming to be the PIC, is operative. An independent experiment could be carried out, ensuring that the maturational level sufficient to make such a constraint visible has been reached. It might look much like the study described in Otsu (1981). Once again, what would be tested by the experimental study suggested in the text is whether the child has or has not "learned" that WH-interpretation is an exception to the PIC, and hence that we may infer that the grammar allowing it to be an exception has or has not been constructed by the hypothesis mechanism.

[50] The interpretation we are assuming for the sentences in (76) is of course that in (i)

(i) a. Where (is) the small trailer (that) he should pull?
b. Where (is) the other one (that) Joe should drive?

The absence of the copula in (76), while interesting, is not relevant. It should not, for example, be taken as an indication that the questions should have the interpretation in (ii):

(ii) a. Where should he pull the small trailer?
b. Where should Joe drive the other one?

At the stage in question, strings with this interpretation would have the form of (iii).

(iii) a. Where he should pull the small trailer?
b. Where Joe should drive the other one?

Given the regularity of the patterns observed for these children, it seems fairly clear that to deny that the strings in (76) involve extraction of WH from a clause would result in a significant loss of generalization. It is data such as these, occurring alongside the other two sets of data we consider in the text, which bear on the question of systematicity raised in Chapter 1. The systematicity of some set of data may be confirmed by the occurrence of other data which the analysis of the first set predicts.

[51] Further support for this alternative output comes from WH-questions involving parentheticals, such as those in (77).

(77) a. (i) Who, in your opinion, did John have it in for?
(ii) How, in the name of Heaven, will we survive?
(iii) Which book, of all that you have read, should the company consider for publication?
(iv) Where, given the opportunity to choose, would you settle?
b. (i) *Who did, in your opinion, John have it in for?
(ii) *How will, in the name of heaven, we survive?
(iii) *Which book should, of all you have read, the company consider for publication?
(iv) *Where would, given the opportunity to choose, would you settle?

It is generally accepted that in the derived structure, such parentheticals are daughters of S, or at least, are followed by a constituent (in this case, S) (Emonds 1976). Such analysis would be difficult if in (77a), the fronted element of M were in COMP, as well as the WH element. Further, if the fronted M in WH questions containing parentheticals could occur in COMP, we would expect strings such as (77b) to be well formed, but they are not.

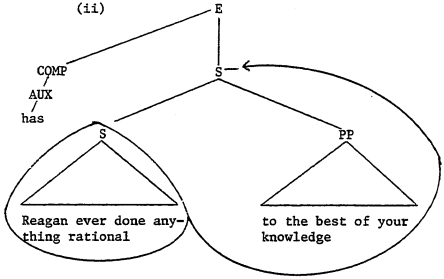
This evidence does raise a potential problem, nonetheless. If we maintain that parentheticals are attached to S, strings such as (i) are also predicted to be well-formed if SAI places AUX into COMP.

(i) a. Was, did you notice, a good set of prints left on the gun?
b. Has, to the best of your knowledge, Reagan ever done anything rational?

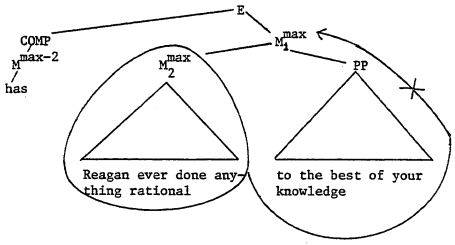
P. Culicover has suggested that such strings are ungrammatical. If they are, they raise an interesting problem. Even if the claim about parentheticals is not just that they must be attached to S, but that whatever follows them must be analyzable as a constituent, the

ungrammaticality of (i) would suggest that S without AUX is not a constituent in fact.

However, there may be an alternate source of the ungrammaticality of (i). In Emonds' (1976) analysis of such parentheticals, it is argued that given the source for these strings shown below in (ii), their appearance sentence internally is accounted for by movement of the leftmost constituent, as the arrows indicate. (Note that the placement of has is determined by the claims made here.)



Within the framework proposed in 3.1 here, the structure in (ii) would be realized as the structure in (iii), if we were to assume that S is the maximal projection of M, and S is E.



Such a structure may be ruled out in principle, since M_1^{max} has no head. However, if M_2^{max} were interpreted as the head of M_1^{max}, the movement indicated by the arrow in (iii) would be prohibited by the Fixed Nucleus Constraint. Thus, sentences like (ib) would be ruled out, if we assume that sentence-internal parentheticals are formed by movement of the left-most constituent.

The details of such an interpretation of Emonds' (1976) analysis would require much more attention before the interpretation could be considered. Nonetheless, its proposal provides a potential source for the ungrammaticality of strings such as those in (i), if they indeed are ungrammatical.

[52] There are two considerations we must deal with here. Since it is the case that different nodes may be involved in Subjacency--$\bar{S}$, S, and NP in English, only $\bar{S}$ and NP in Italian, for example--as we discussed above, the evidence relevant to determining whether a rule is subject to Subjacency (and therefore a movement) will not be the same across languages. Second, we suggested above that a preliminary hypothesis in Italian is that S is also a bounding node for Subjacency. I believe that the scenario I have outlined in the text for the development of a movement rule for WH might have to unfold subsequent to the reformulation of this hypothesis about bounding nodes in a language like Italian.

[53] We can infer this selective attention to data in other areas of development as they are attested in the literature. It is well documented that in their development of plural morphology, children first evidence well formed irregular plurals, but no rule governed ones, e.g., foot/feet occurs, but dog/dogs does not. At a later stage they exhibit well formed regular plurals, and commit errors such as feets. A likely inference, then, for us to draw here, is that data are being "ignored" in the wake of a partially developed system, in which the rule governed mechanism is quite "strong."

Chapter 4

ANOTHER INSTANCE OF THE OPERATION OF THE SUBSTITUTION PRINCIPLE: PARTICLE-SHIFT IN ENGLISH; EVIDENCE FOR AN INTERMEDIATE FORMULATION

4.0. Introductory Remarks[1]

In this chapter we pursue one of the findings of Chapters 2 and 3. In these discussions it was claimed that the hypothesis mechanism will formulate a transformation first as a substitution. This characteristic of the hypothesis mechanism may be referred to as the Substitution Principle. We argued that indeed the output of the mechanism for the formulation of SAI is a substitution. Further we proposed that assuming such a formulation provided an account within a particular framework of the apparent behavior of this rule and its interaction with other rules in the questions we observe in child language. Finally, although the substitution is a formulation that is highly valued by the hypothesis mechanism, when its instantiation results in a misformulation, the misformulated rule is in fact "delearnable," given the interaction of data available to the mechanism and properties of the mechanism itself.

A proposal that makes a general claim about the operation of the hypothesis mechanism, suggesting, as we have, a particular source of "Dissonance," gains more support if evidence for it can be adduced in

more than one case. The goal here is to provide that support. We will be looking at evidence supporting the claim that the rule of particle-shift in English is formulated first in the developing grammar as a substitution. The discussion toward that goal is organized as follows. We will first establish the formulation of the rule in the fully developed adult grammar. Following this, we will treat the major study of the development of verb-particle constructions in the acquisition literature. In section 4.3 we will discuss the shape of the rule in more detail and the evidence that we have for claiming that it indeed has the intermediate shape of a substitution. Finally we treat the learning problem, discussing the evidence required for the "delearning" of the rule, and the availability of such evidence.

4.1. Particle-Shift in the Adult Grammar

The two best known treatments of verb-particle constructions are those of Fraser (1968) and Emonds (1972, 1976). The essential differences between the two analyses are captured in their respective formulations of the rule of particle-shift, which are shown here in (1) and (2)

(1) (Fraser) $V - prt - NP - X \Longrightarrow 1 - 3 - 2 - 4$

(Condition: OBLIG. if 3 = PRO)

(2) (Emonds) $[_{VP}\ X + V - \begin{bmatrix} NP \\ -PRO \end{bmatrix} - [_{PP} P] - Y] \Longrightarrow 1\text{-}3\text{-}2\text{-}\emptyset\text{-}4$ [2]

The arguments favoring Fraser's analysis are generally focused on the constituent status of idiomatic verb-particle sequences, and the need for the two elements to be contiguous in underlying structure if generalizations about their lexical status are to be stated. Thus, the existence of an idiomatic nominal such as "freak-out" or "call-back"[3] are claimed to support the analysis in (1) because in that account the sequence V-Prt is generated by the phrase-structure rules, and the lexical insertion rules, assumed in such an argument to operate in the base, can operate to introduce these expressions into the structure as contiguous elements.

A second argument along these lines is based on nominalizations such as those underlined in (3).[4]

(3) a. John's calling up of the girl surprised me.
    b. *John's calling of the girl up surprised me.

Arguments in Chomsky (1970) led to an analysis in which these derived nominals were base-generated structures, and, hence, lexical insertion of the crucial elements (here calling up) took place in the base and could not be the result of a transformation. In particular, it could not be the result of any sentence rule,because the string John's calling up of the girl was analyzed as a NP. Such an account thus provided for the distinction between (3b) and (4b),

(4) a. John's calling up the girl surprised me.
    b. John's calling the girl up surprised me.

In the framework within which the account of (3) is proposed, a sentence rule such as (1) could account for the alternation in (4) since gerundive nominals such as John's calling up the girl were analyzed as being transformationally derived from sentences.[5]

There are responses to these arguments. To begin, Emonds (1972) argues that not all idiomatic lexical entries must be contiguous elements, citing the underlined idiomatic expressions in (5) as examples.

(5) i. (a) John took his students to task.
(b) *John took to task his students.

ii. (a) His proposal will bring the crisis to a head.
(b) *His proposal will bring to a head the crisis.

iii. (a) John wants to put that car to the test.
(b) *John wants to put to the test that car.

iv. (a) The storekeepers took some students for a ride.
(b) *The storekeepers took for a ride some students.

Further, the argument that an idiomatic expression involving contiguous elements on the surface must originate in the base because of the operation of lexical insertion does not itself hold. It has been proposed that lexical insertion need not be done in the base, and that it may well be carried out on surface structures (Chomsky and Lasnik [1977] and the references cited there).

There is an additional observation to make concerning the distinction between (3) and (4). The phenomenon we see in (3) may be restricted to the class of idiomatic particles and may not extend to the class of directional particles. Both (6) and (7) are grammatical, for example.

(6) Her handing back of the papers without grades was a disappointment

(7) Her handing of the papers back without grades was a disappointment.

Thus, arguments supporting an analysis incorporating (1) over (2) based on the distinction between derived and gerundive nominals do

not hold, since the distinctions on which they are based may not themselves exist, and since the facts themselves do not hold for all particles, as (6) and (7) show.

The arguments that Emonds (1972) adduces in support of (2) are all related to the simplification achieved by analyzing postverbal particles as intransitive prepositions, and, hence, as instances of $[_{PP}$ P]. The success of these arguments then supports (2) because in that formulation of the rule, the $[_{PP}$ P] originates after the direct object NP, and, as Emonds points out, this is the base position for PPs subcategorized by transitive verbs.

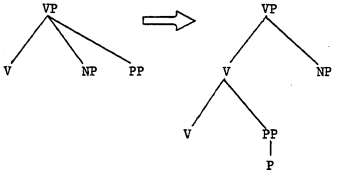
We will assume the analysis including (2) with the following two modifications. First, the condition that the NP be [-PRO] will be removed. It is the case that pronouns can be separated from the verb if these direct object pronouns are contrastively stressed:

(8) a. WHO did they call back?
    b. They called back HER.

The condition that we see operating in verb-particle constructions is an instance of what is probably a general feature of verb-unstressed object pronoun sequences. These pronouns are most likely encliticized to the verb, and therefore they and the verb must have constituent status at surface structure. A condition that is included most plausibly in the set of phonological rules would then block derivations in which the verb and unstressed object pronouns were not adjacent. We will discuss the distinction between the stressed and unstressed pronouns here with respect to the appearance of the rule of particle shift in child language in slightly more detail below.

The second modification is that we will explicitly describe the rule as a Chomsky-adjunction. Thus, the structural change effected by (2) is that in (9).

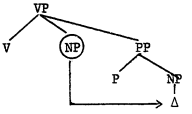
(9)



The resulting V PP sequence thus, would have constituent status in derived structure (being a V), providing a source for the compounds like those we cited above that are constructed from V PRT sequences. Presumably, once it is established that the rule involved is an adjunction, given a rich enough theory of landing sites and rule formulation, no stipulation would have to be included in the rule itself.

The claim made here for the intermediate grammar is that the immature rule, formulated as a substitution--in this case a structure preserving substitution involving the movement of NP--operates as (10) indicates.

(10)



(9) and (10) then reflect an example of the sort of "dissonance" that is possible in the course of acquisition. The claim that (10) reflects the intermediate formulation of particle-shift, and further evidence supporting the claim will be considered in more detail in 4.3. Prior to that discussion, however, we will look briefly at the major account of the development of particle shift in the literature.

4.2. Other Accounts of the Acquisition of Particle-Shift

To my knowledge, only one study treats the development of particle-shift. That is Fischer (1971), The Acquisition of Verb-Particle and Dative Constructions. As the title suggests, however, the study does not treat the development of a rule, but traces the appearance--reflected in the comprehension and production--of these constructions. In this detailed and interesting work, Fischer discusses the analyses of these constructions in the adult grammar, considers the potential interaction of linguistic theory and language acquisition, and presents her findings in a number of experiments dealing with the constructions in child language. The experiments treating the verb-particle construction were of two types; imitation and judgment. In the first type of experiment children were asked to imitate pairs of strings read to them. In the second, they were read the same pair of strings and asked which of the pair sounded better. Justification for the first type of experiment comes from Ervin (1964) and Slobin and Welsh (1969), where it was found that in general "children can repeat sentences only in terms of their grammar" (Fischer 1971, p. 100). The second experiment type was an attempt to test children's intuitions.

In earlier experiments, Fischer (1970) found that children will prefer the last utterance they hear, all things being equal. She reasoned, therefore, that if children's grammars differentiated between grammatical and ungrammatical strings, this preference for recency would be overcome. These two experiments were carried out for pairs such as (11) and (12); Fischer's examples.

(11) a. The girl is calling up her Daddy.
b. The girl is calling her Daddy up.

(12) a. The girl is calling him up.
b. The girl is calling up him.

Fischer found that children (at age 4) had command of the verb-particle construction in sentences such as (11) with full NPs. The data revealed no significant preference for either (11a) or (11b).[6] In sentences such as (12), with pronominal NPs, the children overwhelmingly preferred the order in (12a), consistent with adult judgments. This preference was apparent even when the pronoun was stressed. In this latter case, the children's judgments were at odds with adult judgments as our discussion of (8), above, reveals.

Fischer proposed two alternative accounts for the results obtained with contrastively stressed pronouns. The first is that there is a pronoun output condition applying to unstressed pronouns in the adult grammar, and children do not distinguish between stressed and unstressed pronouns, hence the condition cannot operate in the children's grammar. Her second proposal is that there is a transformational condition in the adult grammar (analagous to the one we dispensed with in (2)) and that it applies in the child's grammar, but prior to the operation of stress placement rules.

There is some support for the claim that children fail to distinguish stressed and unstressed pronouns, and a plausible account of the source for this failure. Baker (1979) proposes that children assign clitic status to object pronouns. This cliticization process is, further, most probably a strong one. Hence, it may well be that data about contrastively stressed pronouns are "ignored," in much the same way that data are ignored in the case we discuss above in Chapter 3. Thus, analyses dependent on the distinction between stressed and unstressed pronouns are not available to the mechanism as hypotheses until the cliticization process is "tempered" by a possible reanalysis of constrastively stressed pronouns as lexical NPs.

Nevertheless, we cannot rule out the possibility that the children's responses to sentences with contrastively stressed pronouns in Fischer's experiment is, in fact, experimental artifact. Because contrastive stress is so much a function of paralinguistic phenomena (such a discourse context, for example, as the necessity of our including (8a) above, emphasizes), children might well ignore stress in strings such as (8b) when they are isolated from discourse and, hence, find them ungrammatical.[7] We need to be careful, therefore, of the inferences about the operation of the mechanism that we draw from these data, even given the plausibility of an account based on the clitic status of the object pronouns.

It is Fischer's discussion of the potential significance for linguistic theory of the facts about children's experimental performance with stressed and unstressed pronouns that is of most interest to us. It reveals again the assumption that pervaded the work in

developmental psycholinguistics at the time, and of which we spoke in some detail in Chapter 1. Consider the following:

> In fact, these findings provide evidence for the theory of stress placement advocated by Chomsky-Halle (1968), which says that the stress rules operate on the surface structure generated by the syntactic component of the grammar. If the syntactic component is not yet fully formed, it is hardly easy for the child to use it for stress placement. (Fischer 1971, p. 97)

Here we see reflected a view of the model of grammar as a dynamic model of acquisition. In that view children's grammars begin as a set of phrase structure rules to which transformations are added as the grammar matures. Surface structure, in such a model, represents the "last step" in acquisition. As we note in Chapter 1, this is not the role linguistic theory assigns to a model of grammar. It reflects misinterpretation of the claims made by the theory. There are no a priori grounds precluding the assumption that the very earliest intermediate grammar (the most "immature" grammar) includes transformations, and there is no theoretical justification <u>for</u> claiming that a model of grammar reflects the chronology of language development. Its inappropriateness notwithstanding, the assumption was clearly a strong one, reflected even in theory-based works such as Fischer's.

The results of Fischer's study of the verb-particle construction allow us, with minor reservations concerning the status of contrastively stressed object pronouns, to infer essentially that children of about 4 years of age do control this construction. They have a rule of particle-shift. The claims argued for in this chapter are distinct from the question of whether or not the rule exists, however. We assume the existence of the rule and question explicitly its shape.

In Chapter 3 it was argued that the initial hypothesis for the formulation of a transformation in the developing grammar favored by the principles that characterize the hypothesis mechanism is a substitution. We turn now to the claim that particle-shift provides an example of the operation of the hypothesis mechanism in this way.

4.3. The "Immature" Rule of Particle-Shift

4.3.1. Formulation of the Rule

Before we consider the evidence for claiming that (10) is the correct output to infer for the rule of particle-shift in the intermediate grammar, we need to look more closely at the formulation of the rule itself, as well as its output, in consideration of some of the issues raised.

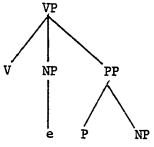
(13) shows what is claimed to be the immature formulation of the rule in the intermediate grammar.

(13) $[_{VP}X + V - NP - [_{PP}P - [_{NP}e]] - Y] \Longrightarrow 1 - \emptyset - 3 - 2 - 5$

One assumption of the general theoretical framework we are assuming is that movement rules leave a trace. The trace of NP movement is an anaphor, and is bound to the moved NP subject to the general conditions on the binding of anaphors (Chomsky 1976, 1980a, 1980b, and the references cited in these works). Of the conditions, the most general is that an anaphor may not c-command its antecedent. (13) violates this condition, since in a structure such as (14), which is a more detailed version of (10), the trace of the moved NP would c-command

its antecedent.[8]

(14)



A reasonable question to ask is why we claim that the mechanism would entertain a rule such as (13) given its violation of the condition on the binding of traces, a condition which, as we shall see, plays a role in children's interpretation of structures related to our discussion of particle-shift--in particular, their interpretation of relative clauses. There is, in fact, a fairly plausible response to the question. We may well be seeing an instance of competition between two principles in the operation of the hypothesis mechanism. During the course of grammatical development, given the data that are input to the mechanism, two (or perhaps more) principles may conflict in the hypothesization of some part of the grammar. When this occurs, one of the principles may be overridden temporarily. We can envision a sequence of events that would culminate in such an override.

In Chapter 2 it was claimed that one "strategy" of the hypothesis mechanism for dealing with new lexical items was to assign them to major categories. The analog of such a strategy is applicable in the parsing of strings by the mechanism. A likely preliminary to the interpretation of a string and the determination of its syntax by the

mechanism is the assignment of constituent structure. Thus, input to the mechanism will be the analysis of P NP sequences as instances of PP. Formulation of the rule of particle-shift as in (13) would allow such an analysis to be maintained by the hypothesis mechanism. We shall see below that in fact the mechanism does analyze the PRT NP sequences in question as PP. Formulation of a rule such as (13) would be favored then by such a parsing strategy which we impute to the mechanism, as well as by the principle favoring substitutions. Further, the relationship between the trace and the moved NP in the derived structure is of no crucial relevance in any other rule, and there would be no information apart from this relationship between the trace and the moved NP about where the trace should be. Although in principle, such information should be sufficient, the parsing information and the substitution principle may temporarily limit the attention the mechanism would ordinarily pay to the relationship between the moved NP and its trace here.[9] Within this scenario there is an important observation to be made. A formulation such as (13) with the output in (14) should be a fairly volatile intermediate grammar, because it does involve a conflict of two principles included in the makeup of the mechanism. We would expect, then, that a grammar generating (14) is fairly short-lived. We will see that evidence of the type required by the "delearnability" requirement is available, also favoring reformulation of the rule.

4.3.2. Evidence for the Formulation of Particle-Shift as a Substitution

Evidence that (13) is the rule we should infer for the intermediate grammar comes from the observation that children do indeed treat the PRT NP sequence as instances of PP. Our evidence is not based on tests involving constituent analysis in the usual sense. That is, we have no data revealing children's intuitions about sentences such as those in (15).

(15) a. Away what did we hide?
b. Off what did we turn?

Analysis of PRT NP sequences as PP by the grammar would entail the grammar's providing structural descriptions for (15), and, hence, children with such grammars would find the sentences grammatical.[10]

The evidence that we do have for the interpretation of PRT NP sequences as PP comes from the interaction of these structures in the interpretation of relative clauses in the developing grammars. To begin, we review briefly some recent findings reported about the interpretation of relative clauses in the course of acquisition.

Of particular interest to us are the findings of Solan and Roeper (1978) and Goodluck (1978). In these studies it was found that children of about 3 or 4-years-of-age interpreted the missing NPs ("gaps") in the relative clauses of (16) and (17) correctly.

(16) The girl pushed the cow that _____ bit the horse.

(17) The boy hits the girl that _____ jumps over the fence.

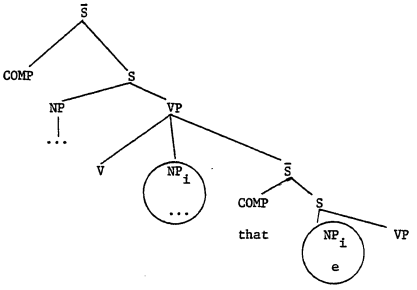
Goodluck further points out that those who would argue that the correct response to (16) is the result of the child's sensitivity to "real world" limitations--girls don't generally bite horses--cannot

so argue in the case of (17). Crucially for us, children of the same ages incorrectly interpreted the controller of the missing NP in sentences such as (18) as being the matrix subject, not the NP <u>the horse</u>.[11]

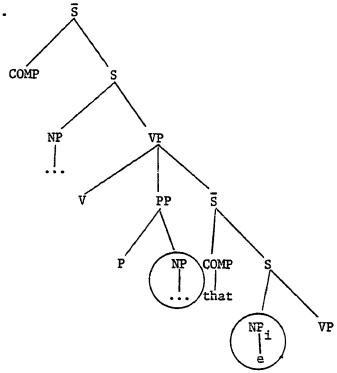
(18) The pig bumps into the horse that _____ jumps over the giraffe.

That is, they are unable to interpret the NP in the PP as containing the controller of the "gap" in the relative clause. If we assume that the hypothesis mechanism makes use of the principle of c-command in the interpretation of these gaps, and that the controller is the nearest c-commanding NP, then the structures in (19), omitting of irrelevant details would account for the interpretation observed in (16), (17), and (18), where the crucial NPs are circled.

(19) a.



(19) b.



The claim made by (19) is that relative clauses in the intermediate grammar are embedded in the VP and not in the NP as they are in the adult grammar.[12]

The datum relevant to the analysis of particle-shift that is proposed here is seen in (20).

(20) The sheep knocks down the rabbit that _____ stands on the lion.

Tavakolian (1978) found children responding to (20) in the same way they responded to (18). That is, the children responded to this sentence as though the NP containing the NP <u>the rabbit</u> were in a prepositional phrase, and thus could not control the missing subject in the relative clause.[13]

The string <u>down the rabbit</u> is a PRT NP sequence. Data such as

the interpretation of (20) support our claim that the developing grammar includes a rule of particle-shift such as (13), resulting in the structure (14). Hence we have independent support for the more general claim that the initial hypothesis for the formulation of a transformation will be a substitution. What remains is for us to consider the evidence required for the mechanism to reformulate (13)--the Learning problem.

4.4. The Learning Problem--Reformulating the Rule of Particle-Shift

While there seems to be a fairly obvious source of evidence that would enable the mechanism to reformulate (13), the situation is slightly complex. The source of the complexity is the number of factors involved in the intermediate grammar that includes (13), all of which need to be re-analyzed by the hypothesis mechanism. The two most critical for us are the PP status of the PRT NP sequences and the placement of the embedded relative clause in (19) as a daughter of VP, rather than of NP. The reason these are critical is that the most obvious type of evidence might allow the reformulation of either without forcing the reformulation of the other.

The evidence of which we are speaking is the contextual interpretation of relative clauses, presumably available to children, and, thus, to the mechanism. The linguistic environment should provide children with the correct interpretation of relative clauses such as those in (21)

(21) a. Daddy bumped into <u>the door</u> that _____ was open.
b. Jason crashed into <u>the pole</u> that _____ held up your Tinkertoy castle.

c. Harry put away the book that _____ was on the table.
d. Suzie snatched up the pair that _____ would have won this game of Fish.

All of (21) would provide children with the evidence that the underlined NPs were indeed controllers of the gaps in the relative clauses. In particular, (21c) and (21d) would provide them with this information about the NPs following particles. In the case of these latter two sentences, would such information "trigger" a reformulation of (13) so that PRT NP sequences were no longer analyzed as PP, or would it, together with the information in (21a) and (21b), cause a re-analysis of the placement of relative clauses?

There do seem to be ways that would allow us to determine which path was being taken. If the mechanism were to re-analyze the structural analysis of relative clauses alone, leaving (13) unchanged, we might expect results in a test of children's judgments of strings like those in (15) that would indicate their interpretation of the PRT NP sequences as constituents. On the other hand, if it were (13) that underwent reformulation, without a change in the analysis of relative clauses, we would expect to find a distinction in children's interpretation of sentences like those in (18) and (20). In (18) they would continue to interpret the subject of the matrix clause as the controller, but in (20), would be able to interpret the NP the rabbit as the controller. Such questions about the developing grammar have not, to my knowledge, been asked, so there are no data in the literature that would tell us which is the path taken by the mechanism. Nonetheless the framework is here for such data to be found.

There is, of course, a third possibility, and that is that the

mechanism re-analyzes the grammar with respect to both aspects simultaneously, even though one feature relates to the statement of the phrase-structure rules and the other to the formulation of a transformation. We claimed above that a grammar with (13) is not likely to be a stable one, given that its existence involves a conflict of two principles of UG. Thus it is possible that when the mechanism is "forced" to examine the structural position of relative clauses, the relationship of (13) to this re-analysis, involving the connection between PRT NP and PP, might in turn force the mechanism to re-examine that rule. Thus re-analysis of (13) would involve both direct positive evidence, of the type that is available in the contextual interpretation of sentences like (21) which are likely to occur in the child's linguistic environment, and indirect evidence, when the mechanism is forced to re-examine an independently unstable grammar. It is, in fact, to some degree, this volatile quality of the grammar itself, as a result of the conflicting principles that may allow the mechanism to abandon a rule such as (13) which is by itself highly valued as an initial hypothesis.

4.5 Concluding Remarks

In the course of this chapter we set out to reveal an additional instance of the operation of the Substitution Principle in the construction of intermediate grammars by the hypothesis mechanism. We saw that there is evidence supporting our inference of an intermediate grammar in which the rule of particle-shift in English is formulated as a substitution. That evidence comes from the interpretation by

children of empty subjects in certain relative clause structures and, in turn, depends crucially on another feature we argue for in the intermediate grammar--the embedded position of the relative clause. Thus this chapter provides further support for proposing that the Substitution Principle, interacting with the data available to the hypothesis mechanism, is a source of Dissonance during the course of acquisition.

In the discussion of the Learning problem, however, a number of other issues were raised. Among these, an important issue concerns the possibility that certain intermediate grammars are more volatile, more subject to reformulation, than others. In fact, this volatile character may account for the fairly elusive quality of the observable evidence in child language which supports our inference of such grammars. The proposed source of the volatile character of the intermediate grammar including the rule in (13) and the structures in (19) which we discuss in this chapter is its dependence on conflicting principles of UG. Proposing such a conflict, clearly, we are making an assertion, if implicit, that our inferences of intermediate grammars may include conflicting principles.[14] Such an assertion broadens the class of possible inferences, and hence of grammars that we can infer. Therefore, we need to examine the quality of volatility and in particular we need to set limits on the way in which principles of UG may interact in the operation of the mechanism. In the following chapter, together with other, related issues we have raised in this and the preceding chapters, we will pursue the problems of setting

limits on how certain principles of UG may interact, and of defining the notion of volatility.

Notes to Chapter 4

[1] For ease of reference, we will, in this chapter, return to the formalism more traditional to the framework of the general theory we are assuming. No claim is to be inferred from its use here, however. Thus, we will use $\bar{S}$, S, NP, VP, and PP in the relevant structures, as opposed to E, M^{max}, N^{max}, V^{max}, and P^{max}.

[2] The VP brackets in (2) are not included in the original formulation of the rule in either Emonds (1972) or (1976). They are added here.

[3] The compound call-back comes from strings such as the following:

They called Barbara back for a second interview.
They called back Barbara for a second interview.

[4] These arguments are reproduced as well in Fischer (1971), who assumes Fraser's formulation as the adult analysis.

[5] The first study so to analyze gerundive nominals was Lees (1960) which is essentially assumed in Chomsky (1970). But see the response to this argument below.

[6] Fischer was concerned with preference here because she felt that it would provide an indication of which position was more "basic" for the PRT, and hence would be the likely candidate for the underlying representation. However, in fact such information wouldn't tell us much about the correct formulation of the rule. Ideally, it will be the theory and the interaction of other rules that give us this information. The selection of (2) as the assumed analysis in the work here, for example, is supported by the generalizations that can be made by claiming that PRT is an intransitive PP, not by claiming that one or the other order of NP and PRT is more basic.

It has often been the case that prevalence of some order of elements in child language was interpreted as support for the claim that this order was the underlying order. I believe that such claims were possible because of the assumption about the relationship of a model of acquisition which we discussed in Chapter 1, and consider again below.

[7] It is interesting in this regard that Solan (1978) found that children around 3-4 years-of-age do have control of contrastive stress. His experimental paradigm crucially did include the sort of discourse context one might want to argue is necessary. Children were required

to act out sentences such as the following with toy animals,

(i) The lion hit the camel, and then HE hit the elephant.
(ii) The lion hit the camel, and then he hit the elephant.

Solan also reports that positive results concerning children's control of contrastive stress were found in production experiments: Hornby and Hass (1970), and Atkinson-King (1973). Tavakolian (1974) is also reported to have found that for the most part children of 4 years understand contrastive stress as a marker of focus.

[8] I am indebted to Michael Rochemont for pointing this problem out to me.

[9] Alternatively, J. Emonds (personal communication) suggests that it may be the case that rightward movement of NP in general leaves no trace. Such a general property of this type of rule would also account for the apparent violation of the c-command condition on trace binding by stylistic-inversion. Note that such a property would be included in UG, and hence would "allow" the hypothesis mechanism to formulate particle-shift to operate as (10) and (13) reflect.

[10] A fruitful experiment, in fact, would be to test the intuitions about sentences such as (15) in children who exhibit the grammars we are about to describe. Results indicating that children find (15) grammatical would lend further support to the grammar being claimed. Of course such an experiment would entail the claim that the children's grammar at this stage would assign a structural description to such strings with fronted PPs. This claim itself requires testing.

[11] This datum was reported in Sheldon (1974), and is cited in Goodluck (1978).

[12] A structure like (19) is considered by Goodluck (1978), but is rejected on the basis of children's interpretation of sentences such as (i).

(i) The boy was hit by the girl that _____ jumped over the fence.

Children correctly interpret the NP <u>the girl</u> as the controller of the missing NP in the relative clause. I don't believe that (19) should be rejected solely on the basis of evidence such as the interpretation by children of (i) for a number of reasons. Among these, the most significant reason is that assuming the structure in (19) allows us to maintain a consistent generalization about the interpretation of gaps in other obligatory control environments.

It was found in a number of cases that children seem to overgeneralize object control from complements of perception verbs to complements of other verbs as well, thus object control was found for (iib) as well as (iia) (Goodluck and Roeper 1978). This finding

supporting the claim that these are embedded in the VP.

(ii) a. John saw Bill carrying a basket.
b. John hit Bill carrying a basket.

Further, children were also found to interpret the object as controller in infinitive in order to clauses (albeit, reduced ones) overgeneralizing here as well; making the NP Pluto the controller.

(iii) a. Daisy hits Pluto (in order to) to put on the watch.

However, in all cases this overgeneralization was found to be blocked in structures such as (iv), in which the potential controller is inside a PP, and, therefore, given the conditions on control and the placement of the embedded clause in the VP is not available as the controller.

(iv) Daisy stands near Pluto to do a somersault.

Claiming that these complements are embedded in the VP, then allows a strong and consistent generalization about the interpretation of gaps in obligatory control environments when a potential controller is apparently "unavailable." This generalization extends to the distinction in interpretation between (16-17) and (18), if we assume (19) for these as well. If we abandon (19), we are left without an explanation for this distinction. In fact, Goodluck is forced to suggest that children's interpretation of (18) may be the result of a pragmatic overload--they must manipulate three animals. However, the giraffe in (18) doesn't play a role that is in any substantive way distinct from the fence in (i). Such an explanation is therefore weak.

A second reason for not rejecting (19) solely on the basis of children's correct interpretation of strings such as (i) is that such strings may not be relevant to the issue. That is, it is entirely possible that the passive by-phrase in (i) should not be analyzed as a PP. A number of scholars have proposed arguments supporting the claim that the analysis of passive includes a lexical rule, sensitive to grammatical (or thematic) relations (Wasow 1977; Anderson 1977; Bresnan 1976, for example). Further, Roeper, Lapointe, Bing and Tavakolian (1981) propose that for rules such as passive, the child's first hypothesis is a lexical rule. If this is the case, then we might be in a position to argue that the by-phrase at some point is not analyzed as a PP by the parser. The specifics of such an argument as well as the character of evidence that would support it need working out. Nonetheless, even in its form here as a speculative proposal, it, together with the foregoing discussion about the interpretation of obligatory control complement subjects provides support for our not rejecting (19) as an intermediate structure for relative clauses solely on the basis of children's reported interpretation of sentences such as (i).

Another interesting question regarding the postulation of (19)

remains. That question is why the hypothesis mechanism should entertain such a structure instead of embedding relative clauses in NP. In principle the theory of the base we are assuming to be part of the hypothesis mechanism (the general framework of $\bar{X}$ syntax) does not restrict clausal recursion to any particular major category. The claim that the initial hypothesis of the mechanism will always be the most restrictive one (cf. the discussion in section 3.2, and note 42 in that chapter) may provide the outline for an answer. Clausal recursion might thus be hypothesized for a category only after the mechanism has evidence that some lexical item of which the category is a projection can itself be subcategorized for a clausal complement. Then it may be the case that such evidence is more readily available in the data children have in the case of verbs than it is in the case of nouns. That is, children may be more likely to hear strings such as (i) than strings such as (ii), although this is clearly an empirical question, the role of the relative availability of certain data in acquisition has yet to be determined.

(i) a. I hope that you haven't written your valentine on the wall.
b. It seems that someone needs more French Toast.

(ii) a. The possibility that you wrote your valentine on the wall upsets me.
b. Daddy's insistance that you need an Apple II-plus amuses me.

[13] It is of interest that the interpretation of (20) in Tavakolian's work played a role in her proposal of the Conjoined Clause hypothesis, which we discussed in Chapter 1. Here we see an alternative account of that interpretation which does not require the proposal of a principle such as the Conjoined Clause hypothesis.

[14] But see note 9 for an alternative account of the apparent violation of the c-command condition, which is one of the principles involved in the potential conflict here.

Chapter 5

ERRORS, VOLATILITY, INFERENCES ABOUT THE HYPOTHESIS MECHANISM AND RELATED ISSUES

5.0 Introductory Remarks

This chapter is offered as a discussion of what we have accomplished in the foregoing work. As such, it is divided into two sections. In the first, we summarize some of the findings, but do so in the context of a discussion of the relevant issues these findings raise. Section two offers a look at what might at first glance be considered two "negative" findings with respect to one of the principles we claim to be part of the hypothesis mechanism. In the course of the discussion, however, we defend the position that these are not negative results, in fact, and consider how they might be incorporated into our picture of the operation of the mechanism.

5.1. Errors, the Mechanism, and the Character of Intermediate Grammars

In the following pages we will consider some of the general issues that have developed in the course of this study. The purpose here is to provide a summary of the findings and to highlight what we consider to be the most important of these for further work in both the areas of language acquisition involving observation of child language, and of linguistic theory. But I also wish to highlight the

questions raised by some of these issues and provide an outline of what might be included in answers to them.

We can begin by looking at the errors we spoke of in Chapter 1. There we established that the investigation would be limited to looking at systematic errors in child language, and that these errors would be assumed to indicate a grammar that deviated from the adult, fully developed grammar. A goal was to determine the ways in which a restrictive theory of possible grammars (like the theory of UG), which is part of the hypothesis mechanism interacting with the data available to the hypothesis mechanism, would allow such errors. Further we sought to propose ways in which this interaction would limit the types of errors we might observe in child language.

In fact, the word error is used throughout the work with two meanings. In its first sense, it refers to the errors we can see in the language behavior of children. Some of these are directly observable in speech; the absence of Subject-Auxiliary Inversion in all questions, then in just WH-questions, for example, or the repetition in distinct parts of the string of pieces of inflectional verb (=AUX) morphology. Other errors we see indirectly, as a result of the interaction of the mistake with another rule in the grammar. Such was the case with particle movement, where the error surfaced--became observable--only as a result of the interpretation of relative clauses in a structure that was the output of this rule. In this latter type of error, the second meaning of the word surfaces more clearly. This second interpretation relates to the structural descriptions of child language that we are inferring. In this case we are speaking of

errors that are possible in the formulation of the grammar itself. The moment a grammar is inferred on the basis of an error that is observable on the surface, the inference of this grammar makes the implicit claim that the error in formulation that it entails is one that is possible on the basis of the interaction of the hypothesis mechanism with the data available to it. That is, by claiming that an error observable in child language is itself the result of an error in the formulation of some part of the grammar, we are making the claim that somehow the hypothesis mechanism allows--if not provides for--the particular misformulation. Thus, the systematic ambiguity in our definition of the notion error allows us to connect data from child language with Principles of Universal Grammar. The connection is made by our proposals for fragments of intermediate grammars which deviate systematically from the corresponding fragments in fully developed grammars.

One goal of this study then has been to propose and justify analyses, making the argument that indeed the principles which are part of the hypothesis mechanism can interact with available data to result in grammars that are distinct from the fully developed adult grammar. We have seen two examples of this interaction which themselves suggest the "types" of errors we can predict in the course of language development. In one case, we propose an analysis involving a parameter along which we claim grammars can be constructed. The parameter is the placement of COMP with respect to TENSE. The initial interpretation of this parameter by the mechanism was shown to interact with two principles in the theory of grammars--the Propositional Island

Condition of Chomsky (1977), and the Fixed Nucleus Condition of Schwartz (1973)--and what we claimed to be a substantive universal about the base, to be incorporated into the theory--that M(ODAL) is the head of S. The result of the interaction was an intermediate grammar, G II, in which the formulation of any rule moving AUX (which would thus be a movement of some level M^{max-n}) was blocked, and all WH-interpretation was clause-bound.

In the second case, we proposed a principle, the Substitution Principle, favoring the formulation of transformations as substitutions, ceteris paribus. In at least two instances--the formulation of SAI in the intermediate grammar G III where a rule moving M^{max-n} first becomes possible, and in the formulation of particle movement--we saw that the hypothesis mechanism does, given certain conditions, formulate a rule that is an adjunction in a fully developed grammar, first as a substitution.

These two cases themselves suggest that the notion "type of error," which we originally sought to be able to predict, might itself be misguided. In the case of G II, one could say that the error was the absence of a transformation--no SAI--or that it was the misformulation of the base--COMP being generated in the domain of M. But we would be missing the important issue, which is that there is a succession of grammars in the course of language development, and therefore what we need to look at are the conditions determining the shape of these grammars--or at least the parts of them we look at. Even in the case of the misformulated substitutions we could not look just at the rule itself, but we needed to look at its interaction with

the rest of the grammar, as well as with the hypothesis mechanism.

With this view of language development we are able as well to isolate two other areas in which to extract generalizations. The first we will think of as generalizations about the behavior of the mechanism and the inferences we can make about this behavior. The second we will describe as generalizations about the character of intermediate grammars. In the study here, we have seen that in the operation of the mechanism there can be a conflict of principles--exemplified in the case of the formulation of particle movement as a substitution, moving NP to the complement position in a PP from its position as right sister to V. The conflict of principles in this case involved c-command and the primacy of substitution. We have also seen that the mechanism can be argued to "ignore" data when it has constructed a grammar that cannot assign structural descriptions to the crucial data. The transition from G I to G II in the development of SAI provided us with an example of this behavior. There we saw that the mechanism ignored evidence about the alternation of AUX in preverbal and sentence-initial position, because it could not assign a structural description to strings with fronted AUX given the impossibility of such a movement in the grammar with COMP in the domain of TENSE, as we defined it. We also claimed that the mechanism ignored data relevant to the reformulation of the grammar with respect to COMP at this time; data involving WH-interpretation in English with respect to the PIC. We proposed that the reason for this might be the immaturity of the mechanism, and cited other cases where such an explanation might be appropriate, as opposed to the conclusion that

principles of UG do not participate in language development and are not part of the biological endowment that the mechanism brings to the task. Finally, we maintained that the mechanism will only construct grammars that can be reformulated on the basis of the data that are available to it. In general these data are assumed to be those which do not include explicit direct evidence concerning the ungrammaticality of any given string in the language. This we refer to as the Delearnability Condition.

This condition is stated as a constraint on the types of grammars we may infer as outputs of the mechanism. It can also serve as a constraint in our inferences about the way the mechanism itself might go about constructing a grammar. For example, we suggested ways in which the requirement justifies our proposing the way the mechanism might go about determining the bounding nodes for subjacency in the language whose grammar it is constructing. Further, such a requirement supports our inclusion of something like the Substitution Principle in the hypothesis mechanism. Finally, we can see the requirement reflected in our suggestion in Chapter 3 that the hypothesis mechanism initially assumes that Surface Structure and the input to Logical Form are isomorphic, and that no rules "reconstructing" Surface Structure for the interpretive rules are posited without positive evidence. Thus, we redefine a requirement that has been proposed to be a constraint on possible grammars in Linguistic Theory (cf. Baker 1979) as a constraint on our inferences about the grammars that are outputs of the hypothesis mechanism, and, by extension, on our inferences about the operation of the mechanism in the construction of

these grammars.

In the second area, the character of the intermediate grammar, there are two very general observations we can make about the grammars that are not properly forthcoming from the discussion of any of the issues relating strictly to Delearnability. We have seen that certain grammars are highly volatile, and hence, short-lived, while others are more stable, and reflexes of them in child language exist for some time. A highly stable grammar is G III, involving the misformulation of SAI and it is the case that errors reflecting it occur for some time in child language.[1] G I is a grammar which we have claimed is a volatile one, and its construction entails the mechanism ignoring data. Our claim about its volatility gains support from the elusive character of the data in the literature which contribute to its justification, insofar as this elusiveness of data can be interpreted as a function of the short life of a grammar. The source of the volatility of the grammar including the (mis-)formulation of the rule of Particle-Shift we claim is the conflict of principles that it entails. We saw that the data supporting our inference of the grammar is extremely limited. However, in both cases, the elusiveness or paucity of data is not itself sufficient to justify our labelling the intermediate grammars as volatile. Having established the existence of the grammars, we need to determine the length of their "lives" in the course of language development in order to determine that they indeed are relatively short-lived. Since no accounts have explicitly proposed either grammar, we need to return to the arena of observation, look for evidence of the grammars--sentence-initial co-occurrence of

WH and AUX in the absence of independent evidence for movement in the case of G I and interpretation of Prt-NP sequences as PP in the case of the grammar including the misformulation of Particle-Shift--and establish the boundaries of their existence.

Being able to outline ways in which we may identify grammars as volatile has to be seen as a positive result itself. This is because such a label should make us more careful about what we accept as possible descriptions. Since we want to talk about the behavior of the mechanism--how grammars are formulated in the course of language development--we do at some point need to posit structural descriptions of certain phenomena. The more limits we have on what may qualify as plausible descriptions, the better our picture of the mechanism we are claiming constructs them will be. As a result, the interaction between linguistic theory and the observation of child language will have a chance of being more fruitful.

Thus far we have talked about the ambiguity in our definition of the word <u>error</u>, the relationship of the notion "error type" to our goal of being able to talk about the shape that intermediate grammars may have in the course of language development, and the relationship of the grammars we claim are possible to the behavior of the mechanism. There is one other subject that deserves attention, and that is the notion of systematicity. When we spoke of the "errors" that were observable in child language, we asserted that only those that were <u>systematic</u> would contribute to the justification of the inference of grammars providing structural descriptions for them. We also maintained that confirmation of the systematicity of certain errors would

have to wait for some of the work to be done. I think that it is clearer now why that was the case. Systematicity is not determined only on the basis of the pervasiveness and persistence of an observable error in child language. It relates as well to the presence (or absence) of other phenomena predicted by the grammar inferred on the basis of the original, observed error. Thus, in the case of G I, which generates structures in which no SAI can occur, a consequence of this grammar is the generation of structures in which WH-interpretation will be clause-bound, and there is evidence indicating that indeed WH-interpretation is clause-bound at this stage in child language. As well, the formulation of particle-shift as a substitution predicts that the output of the rule will have a particular structure, namely PP. We saw that this prediction interacted with other parts of the misformulated grammar--the analysis of the embedding of relative clauses--culminating in errors in the interpretation of the "gap," the phonetically null NP, in certain relative clauses. When we see this clustering of properties that can be analyzed as consequences of assuming that the original error was systematic, we have additional confirmation that the original assumption of systematicity was correct. Thus, we do not have to depend on any intuitions we might have about what is or is not systematic in child language. In the course of this kind of work, we can test our assumptions by looking for this clustering of related phenomena in the grammar.

5.2. "Negative" Results

I wish to turn now to what may appear as negative results in two areas, but in fact are positive results of an interesting kind.

5.2.1. The Case of Verb-Second in German

In Roeper (1972) it was observed that children acquiring German as their native language did not make the error of failing to invert the finite verb in WH questions.[2] Thus, for a question such as (1), children were never observed to produce the forms in (2).

(1) Warum spielt er Fussball?
'why plays he football?' 'Why does he play football?'

(2) a.* Warum er spielt Fussball?
b.* Warum er Fussball spielt?

Assuming that the underlying word order in German is SOV, and that there is a rule of verb second, moving the finite verb to second position in root clauses, Roeper proposes the following explanation for the absence of the error:

> The explanation is that Germans must create declaratives and questions with the same verb-second transformation. It has, therefore, much broader scope than the subj-verb inversion transformation in English; hence it is acquired earlier. (p. 41)

If we claim that the error observed in the speech of children acquiring English is the reflection of a grammar in which the inversion rule is misformulated as a substitution, however, such an explanation itself is not sufficient to explain why it might be the case that the rule is not formulated first as a substitution in German. Either our generalization about the formulation of such rules as substitutions being favored is wrong, or there is some clue in

German, to which the mechanism has access, that will override the principle, and allow the mechanism to posit the correct formulation of the rule immediately. That is, we need to look for some difference in the grammars of English and German from which it would follow that finite verb movement in German cannot be formulated as a substitution.

A possibility is that the rule in German applies exactly when COMP is filled.[3] This possibility assumes first that the rules of Verb-Second and Inversion in questions are, in fact, the same rule in German, and that in that language, as in English, the rule involves an adjunction.[4] Verb-Second applies in root clauses and we observe all of the following strings; none of which are claimed to have any exceptional illocutionary readings (involving FOCUS, for example):

(3) a. Gestern hat Hans das Buch dem Mann gegeben.
'yesterday-has-Hans-the Book-the man-given.'
b. Hans hat das Buch dem Mann gestern gegeben.
c. Das Buch hat Hans dem Mann gestern gegeben.
d. Dem Mann hat Hans das Buch gestern gegeben.

In this sense, we can say that the rule in German analogous to SAI in English is "more general." But it is a feature of the grammar of the rule which makes it more general--that it is exclusively an adjunction, triggered by a filled COMP.

The immediate response to such a proposal is to point out that direct YES-NO questions in German have the same shape as such questions in English: the fronted finite form is in sentence-initial position with no constituent preceding it:

(4) Kommst er heute?
'comes he today'
'Is he coming today?'

Thus there appears to be no sentence-initial constituent "triggering" the rule. Most traditional Transformational treatments of SAI for English in the literature, however, do assume some underlying abstract marker, Q (Baker 1970), or the feature +WH, for example. Until recently, however, such features never resulted in an analysis of the node to which they were attached as filled. The feature +WH in COMP, for example, was involved crucially in the argument supporting the analysis of WH-fronting as a structure preserving rule (cf. Emonds 1976). But, the question remains: if even the presence of such a feature in the structure allows the node COMP to be analyzed as empty in English, why might it not be analyzed as empty in German?

The answer in fact may be that it is not necessarily analyzed as empty; or at least that is not crucial to the mechanism acquiring German that it be empty in this case. If Verb-Second and Inversion are the same rule in German, and if the Verb-Second cases are learned first, then Inversion is acquired as a "special case" of Verb-Second. The mechanism has stable evidence about the "landing site" for the fronted finite verb form, based on its formulation of Verb-Second, so the availabiliby of formulation as a substitution is not required to resolve any problem such as that one. Once there is stable evidence supporting formulation of the rule as an adjunction, the mechanism has no reason to reformulate the rule as a substitution in a related, if slightly different, environment.

It is this aspect of the general nature of the rule that insures its initial formulation as an adjunction in the developing grammar of German. That is, in its general case--in a basic and early acquired

rule, Verb-Second--preposing of the inflected verb form occurs exactly in the presence of a filled COMP, and is formulated securely as an adjunction. There is no compelling reason for the mechanism to reformulate it in the special case of YES-NO questions. Hence the case of WH questions presents no special problem to the mechanism.

An argument could be made that in English as well, there are cases of SAI which involve a filled COMP. In support of such an argument, one might cite examples such as the following:

(5) a. So ambitious was she that she destroyed even those who would support her.
b. Not in a million years did he imagine his own gun would do him in.
c. Only with great care should you stalk a hungry rhinocerous.

But, unlike the case in (3) above, all such cases in English, involving Adverb preposing, or negated constituent preposing do entail the focusing of the preposed constituent that co-occurs with the fronted AUX. Thus such strings are (at least pragmatically) marked, and may not contribute to the early development of the grammar. We noted in Chapter 2 that echo questions--another marked structure--do not play a central role in the analysis of WH constructions by the mechanism. While this discussion does involve speculation about the role data can play in the formulation of intermediate grammars--speculation of which we are skeptical in general--drawing a distinction between the role strings such as (5) and strings such as (3) may play in the formulation of the grammar seems justifiable.

What we have done here is to suggest a way in which Roeper's notion of generality can be made more precise. He proposed that because the rule in German applied in a wider variety of contexts;

whether the first element was the subject, an adverb, or a WH constituent, for example, it was a more general, and hence simpler, rule than its counterpart in English. What we have proposed is that the operation of the rule in German, adjoining the finite verb either to whatever is in COMP, or to S (M^{max}) in the presence of some element in COMP in all of these contexts prevents its formulation as a substitution in the one context where such a formulation is possible--YES-NO questions. In English, formulation of the analog, SAI, as an adjunction is not indicated by evidence which we maintain would be readily available to the mechanism, and, hence, the principle we claim to be part of the mechanism favors initial formulation of the rule in English as a substitution first.

5.2.2. The Case of Clitics in French

The second case of an apparent negative result comes from French. In Klein (1977) it was proposed that given the correctness of the principle favoring the formulation of transformations as substitutions by the hypothesis mechanism, a "type" of error we should explicitly look for is the misformulation of adjunction rules as substitutions. In the foregoing discussion we have already touched upon the misguidedness of limiting a research program to looking for "types" of errors such as this. Rather, we need to look at the way proposed principles may interact with other principles in the mechanism as well as with data in the environment to construct intermediate grammars, and subsequently to reformulate the ones that are in "Dissonance" with the data that are the projection of fully developed, adult grammar.

If we were, nonetheless, to look for evidence of an adjunction misformulated as a substitution, it was suggested that the grammar of clitic-placement, specifically the rule of le, la, les Movement, provided such an example. In that work, the analysis assumed for the adult grammar was that of Emonds (1975, 1976). The rules in his account that are crucial for us here appear below in (6).

(6) a. $V' \rightarrow \left\{ \begin{array}{c} V' \\ (PRO)-(CL)-TENSE \end{array} \right\} - V$

b. le, la, les Rule

$$X - V' - \begin{bmatrix} PRO \\ +III \\ -REFL \\ \alpha PLUR \\ \beta FEM \\ NP \end{bmatrix} - y \Longrightarrow 1 - \begin{bmatrix} DEF \\ \alpha PLUR \\ \beta FEM \end{bmatrix} + 2 - \emptyset - 4$$

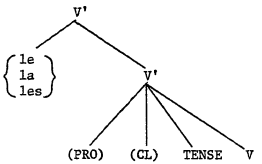
c. Pronominal Clitic Placement:

$$X - [_{V'}[^{\Delta}_{PRO}] - Y] + Z - [_{NP}(\grave{a}) - PRO] - W \Longrightarrow 1 - [^{5}_{[-FEM]}] - 3 - \emptyset - \emptyset - 6$$

In (6a) we see the base rule that expands the verbal complex in Emonds' account. Justification for the base position CL, which accounts for the placement of y and en (the "PRO-PP," or adverbial clitics) by a structure preserving movement, comes from cases of en that cannot be argued to be related to de NP; s'en aller, en avoir marre de NP

('to get out,' 'to be fed up with NP'). The PRO position is justified by the existence of what are referred to as intrinsic reflexives; the se in se laver la tête ('to shampoo one's hair'), for example.[5] It is into this position that the rule in (4c) moves all object pronouns that are not analyzed by the rule in (6b). The rule in (6b), "transforms (postverbal) direct object pronouns (lui, eux, elle, and elles) into the preverbal 'definite article' forms le, la, les" (Emonds 1976, p. 233). While Emonds claims that (6b) is not strictly a movement rule, but rather a combination of a deletion and an insertion of an element that agrees with the deleted element in grammatical number and gender, nothing in his analysis rests crucially on that view of the rule, and that view may be left as is for our purposes here. Further, Edmonds does not speculate about the derived structure of a rule such as (6b) in this work, but in Emonds (1978), he maintains that le, la, les and V' are a constituent. We may infer (7) as the plausible output structure, if we assume (6b) operates as a Chomsky adjunction whether the clitics are introduced by movement or insertion. For clarity, we include the full expansion possibilities available in (6a).

(7)



The proposal in Klein (1977) was that the le, la, les rule would be formulated first in child language as a substitution, most plausibly for the position PRO generated by (6a). If this were the case, the prediction follows that there would, at the time when the rule was misformulated, be no combinations of clitics such as (8).[6]

(8) a. Il le lui donnera. 'He'll give it to him/her.'
b. Je te l'emprunterai. 'I'll lend it to you.'

The data relating precisely to this question have not been available in French. However, there are related data from Spanish, where, although a number of studies differ as to the correct analysis of the source of clitics (Strozer 1976; Rivas 1977), an argument could be made for rules parallel to those here in the case of the substitution for PRO and the le, la, les rule. Thus we have data such as (9).[7]

(9) a. Se lo compraron. 'They bought it for her.'
b. Se los do. 'He gave them to him.'
c. Una muchacha me lo hizo. 'A girl made it for me.'

These strings involve the co-occurrence of the direct and indirect object pronouns in clitic positions analogous to the positions in (8).

Let us suppose, for a moment, that these data can be interpreted as showing that the prediction about a rule such as the le, la, les rule is false--it is not formulated first as a substitution. The mechanism formulates it correctly from the beginning as an adjunction. What interaction with other areas of the grammar might allow the mechanism to override the principle about substitutions and allow the formulation of the rule as an adjunction?

If we return to the French account, we can find evidence that would be relevant to the problem, and available to the mechanism.

The evidence is in the shape of data indicating the constituent status of the V' and adjoined clitic. That is, le, la, les and the following V' form a constituent (V'). We have discussed some of these arguments above, in Chapter 3, where it is proposed that the constituent in question is M^{max-2}. Once again, the support comes from rules which have been observed to analyze the sequence of the clitic and the left-most verb as a constituent. The two clearest cases are the rules of Subject-Clitic Inversion and Auxiliary Deletion (Kayne 1975). Examples of the operation of these two rules are shown in (10) and (11) respectively, which are essentially the same as (27) and (29) in Chapter 3.

(10) a. Me l'as tu donneras? 'Will you give it to me?'
b. Le leur as tu donné? 'Did you give it to them?'

(11) a. Paul m'a bouscoulé et _____ poussé contre Marie.
'Paul jostled me and pushed me into Mary.'

b. Paul l'a insulté et _____ mis à la porte.
'Paul insulted him and threw him out.'

We could say then, that the mechanism has information about the constituent status of the output of the le, la, les rule that is strong enough to override the principle favoring formulation of the rule as a substitution. The kind of evidence that we have here is not available in either the case of Particle-Shift or SAI in English. There are essentially no rules that analyze the output of SAI, and information about the output structure of Particle-Shift is really available only in the form of single lexical items, such as put-on, take-out, and the like. Further, we have claimed that reformulation of the misformulated Particle-Shift rule depends on a fairly intricate interaction of the interpretation of the output of the

misformulated rule and the interpretation of the missing NP subject in relative clauses on the NP in the V PRT NP sequence.

If it is the case, then, that a rule such as the le, la, les rule in French is not misformulated in accordance with the principle favoring formulation as a substitution, we can propose that it is because the mechanism has "secure" information about the output structure of the rule, given the availability of the relevant evidence. Thus, we do not, in fact, have what must be interpreted as a negative result. Even with this account of the why it might be the case that a misformulation is "blocked," I am not prepared to conclude that it in fact is, solely on the basis of the evidence here from Spanish. That is because all of the strings in (9), as well as other relevant strings in the data from which they are taken, are reported as being from children over five years of age. The particular data in (9), in fact are reported to have occurred in the speech of the children of six and nine years. The misformulated substitutions we have isolated in English occur in children between the ages of about three, in the case of SAI, and four, in the case of Particle-Shift. The question about the formulation of the le, la, les rule, therefore, deserves to be posed in an experimental setting with children of comparative ages acquiring French as their native language.

While the thrust of this work does not entail such an experimental investigation, I would like, as a final note, to suggest what it might involve. To begin, an imitation task should be included, given the correctness of the assumptions about the relationship of performance on such a task to the structure that it is correct to infer for the

grammar. But as well, we could design an experiment that would explicitly attempt to elicit these sequences of clitics. Such a design would necessarily involve some attention to the pragmatics involved in the use of clitics in discourse, so that artifact effect does not blur our picture of what might be the right grammar to infer. A short story with pictures, followed by a question requiring an answer in which the use of clitics would be appropriate might elicit the responses we require. A sketch of such appears in (12).

(12) a. (i) Paul a reçu un cadeau d'anniversaire.
(ii) Marie a donné $\left\{\begin{matrix}\text{le}\\ \text{ce}\end{matrix}\right\}$ cadeau à Paul

b. Comment Paul, a-t-il reçu le cadeau?

c. "required" answer: Marie le lui a donné.

relevant "wrong" answer: Marie l'a donné à lui.

It is also possible that an elicited answer might be (13)

(13) Marie lui a donné le cadeau.

While this response would also be relevant to us, it is less likely, given the relationship of other phenomena, such as focus, and presupposition, for example, to the use of clitics. Nonetheless, the optimal design of an experimental paradigm would require us to be able to tease out the factors attributable to these other phemonena.

5.3. Concluding Remarks

What we have done here, then, is raise two examples of what might be considered negative results, and discuss their relevance to the picture we are constructing of the role played by the principles we attribute to the hypothesis mechanism in the course of language

development. The goal in raising them was to make the picture clearer--to use these examples in an examination of how the principles (one in particular) might interact with other principles and with other fragments of the developing grammar. It is hoped that progress toward that goal is apparent, and that the possibility of carrying out the experimental investigations suggested here (as well as elsewhere in the text, cf., Chapter 3) will provide additional progress in that direction. It is my belief that the direction is the right one. What we require for such interaction between linguistic theory and observation of data in child language to be truly fruitful as well, is to continue to go in this direction even as the theory changes. We will be able to adapt to changes that make the theory more correct, because the research paradigm suggested by the work proposed in the text here has a life that is identifiably distinct from a particular model. The results here <u>are</u> within the framework of a particular (if fairly general, given the principles we include) theory, but the work that yields such results may be continued, the questions reformulated, as the theory evolves.

Notes to Chapter 5

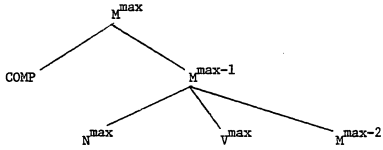
[1] In fact the very stable ones, G III, for example, are reflected in errors that are observable in other domains in which grammar construction arguably is taking place--second language acquisition. The distribution of SAI in the speech of speakers acquiring English as a second language has been widely observed to mirror its distribution in the speech of children during the time we claim G III is justified; SAI occurs in YES-NO questions, but not in WH questions.

[2] Roeper also discusses this problem in Roeper (1973), "Theoretical Implications of Word Order, Topicalization and Inflections in German Language Acquisition."

[3] The discussion that follows is based on the work of a number of scholars who have considered the problem of the rule of Verb-Second in German and Dutch and its relationship to inversion in questions. Included in a list must be the following: denBesten (1977), Koster (1978) and more recently, Thiersch (1978), Safir (1981).

[4] For more detail on how the adjunction is to be formulated in English; whether M^{max-n} is adjoined to COMP or Chomsky-adjoined to M^{max}, see Chapter 3, section 3.1.4. Two other assumptions are implicit in the discussion here. One is that in German when Verb-Second applies, the inflected verb is dominated by M (i.e., verb raising has taken place--cf. Chapter 3), and that given the framework proposed in Chapter 3, the structure of strings prior to Verb-Second is as in (i).

(i)



The question about the placement of COMP with respect to M^{max} remains here as well, since the rule in question must move the head, M^{max-2},

out of its phrase to be well-formed. The second assumption is that in declaratives, the subject NP is placed in COMP. Koster (1978) argues that subjects do appear in COMP, in the shape of subject sentences, and Thiersch (1978) supports the claim that in each of the sentences in the text, taken from Safir (1981), the initial constituent is in COMP. An important question is how the mechanism might determine that the subject in German is in COMP. This is in fact a question raised by K. Wexler. There is a plausible account of the course of such a determination. Once the mechanism determines that there is a rule of Verb-Second which moves M^{max-2}, it can only formulate the rule as movement to COMP in a structure such as (i), given the Fixed Nucleus Constraint. Thus whatever constituent appears to the left of the fronted M, including the subject NP, must be inferred to be in COMP as well.

[5] Emonds attributes these observations originally to Kayne (1970).

[6] (6b) involves as well a rule interchanging nonthird person PRO clitics and le, la, les. This rule is not of direct relevance to us here, however.

[7] These data come from the work of Martha Randeri (1976), who traces the use of clitics in children from the ages of 5-9 years.

BIBLIOGRAPHY

BIBLIOGRAPHY

Akmajian, A., S. Steel and T. Wasow (1979) "The Category AUX in Universal Grammar," Linguistic Inquiry 10, 1-64.

Akmajian, A. and T. Wasow (1974) "The Constituent Structure of VP and AUX and the Position of the Verb BE," Linguistic Analysis 1, No. 3.

Andersen, H. (1973) "Abductive and Deductive Change," Language 49, 765-793.

Anderson, S. R. (1977) "Comments on the Paper by Wasow," in P. W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Andersson, L. (1975) The Form and Functioning of Subordinating Clauses, University of Göteborg Dissertation.

Atkinson-King, K. (1973) Children's Acquisition of Phonological Stress Contrasts, UCLA Working Papers in Phonetics, 25.

Baker, C. L. (1970) "Notes on the Description of English Questions: The Role of an Abstract Question Morpheme," Foundations of Language 6, 2.

Baker, C. L. (1977) "Comments on the Paper by Culicover and Wexler," in P. W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Baker, C. L. (1978) Introduction to Generative-Transformational Syntax, Prentice-Hall, Englewood Cliffs, New Jersey.

Baker, C. L. (1979) "Syntactic Theory and the Projection Problem," Linguistic Theory 10, No. 4.

Baltin, M. R. (1978) Toward a Theory of Movement Rules, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Bellugi, U. (1967) The Ac uisition of Negation, Harvard Doctoral Dissertation.

Bellugi, U. (1965) "The Development of Interrogative Structures in Children's Speech," in K. Riegel, ed., The Development of Language Functions, Ann Arbor, Michigan Language Development Program.

Bellugi, U. and R. Brown (1964) "Three Processes in the Child's Acquisition of Syntax," Harvard Educational Review 34, 2, 133-151.

Besten, H. den (1977) "On the Interaction of Root Transformations and Lexical Deletive Rules," unpublished paper, Instituut voor Algemene Tallweterschap, University of Amsterdam, Netherlands.

Bloom, L. (1973) One Word at a Time: The Use of Single Word Utterances before Syntax. Mouton, The Hague.

Bowerman, M. (1973) Early Syntactic Development, Cambridge University Press, Cambridge.

Braine, M. (1976) Children's First Word Combinations, Monographs of the Society for Research in Child Development, No. 164, 1, No. 1.

Bresnan, J. (1970) "On Complementizers: Toward a Syntactic Theory of Complement Types," Foundations of Language 6, 297-321.

Bresnan, J. (1971) "Sentence Stress and Syntactic Transformations," Language 47, 257-281.

Bresnan, J. (1978) "Realistic Model of Transformational Grammar," in M. Halle, J. Bresnan and G. Miller, eds., Linguistic Theory and Psychological Reality, MIT Press, Cambridge, Massachusetts.

Brown, R. (1968) "The Development of Wh Questions in Child Speech," Journal of Verbal Learning and Verbal Behavior 7, 279-290.

Brown, R. (1973) A First Language: The Early Stages, Harvard University Press, Cambridge, Massachusetts.

Brown, R. and C. Hanlon (1970) "Derivational Complexity in Child Language," in J. R. Haynes, ed., Cognition and the Development of Language, Wiley, New York.

Chomsky, N. (1957) Syntactic Structure, Mouton, The Hague.

Chomsky, N. (1964) "Current Issues in Linguistic Theory," in J. A. Fodor and J. J. Katz, eds., The Structure of Language: Readings in the Philosophy of Language, Prentice-Hall, Englewood Cliffs, New Jersey.

Chomsky, N. (1965) Aspects of the Theory of Syntax, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1972a) "Some Empirical Issues," in S. Peters, ed., Goals in Linguistic Theory, Prentice-Hall, Englewood Cliffs, N.J.

Chomsky, N. (1972b) "Deep Structure, Surface Structure and Semantic Interpretation," in N. Chomsky, ed., Studies on Semantics in Generative Grammar, Mouton, The Hague.

Chomsky, N. (1973) "Conditions on Transformation," in S. R. Anderson and P. Kiparsky, eds., Festschrift for Morris Halle, Holt, Rinehart and Winston, New York.

Chomsky, N. (1977) "On WH-Movement," in P. W. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Chomsky, N. (1979) "Markedness and Core Grammar," unpublished paper, MIT, Cambridge, Massachusetts.

Chomsky, N. (1980a) "Principles and Parameters in Syntactic Theory," mimeographed, MIT; to appear in N. Hornstein and D. Lightfoot, eds., Explanation in Linguistics, Longmans, New York.

Chomsky, N. (1980b) "On Binding," Linguistic Inquiry 11, 1.

Chomsky, N. (1981) Lectures on Government and Binding: The Pisa Lectures, Foris Books, Dordrecht.

Chomsky, N. and M. Halle (1968) The Sound Pattern of English, Harper and Row, New York.

Chomsky, N. and H. Lasnik (1977) "Filters and Control," Linguistic Inquiry 8, 3.

Culicover, P. W. (1976) Syntax, Academic Press, New York.

Culicover, P. W. and K. Wexler (1974) "The Invariance Principle and Universals of Grammar," Social Sciences Working Paper No. 55, University of California, Irvine.

Culicover, P. W. and K. Wexler (1977) "Some Syntactic Implications of a Theory of Language Learnability," in P. W. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Dale, P. S. (1976) Language Development, Structure and Function, Holt, Rinehart and Winston, New York.

Derwing, Bruce (1973) Transformational Grammar as a Theory of Language Acquisition: A Study in the Empirical, Conceptual and Methodological Foundations of Contemporary Linguistic Theory, Cambridge University Press, Cambridge.

DeVilliers, J. and P. De Villiers (1978) Language Acquisition, Harvard University Press, Cambridge, Massachusetts.

Drachmann, G. (1979) "Child Language and Language Chance: A Conjecture and Some Refutations," in J. Fisiak, ed., Recent Developments in Historical Phonology, Mouton, The Hague.

Emonds, J. E. (1970) Root and Structure Preserving Transformations, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Emonds, J. E. (1972) "Evidence that Indirect Object Movement is a Structure Preserving Rule," Foundations of Language 8, 546-561.

Emonds, J. E. (1974) "Arguments for Assigning Tense Meanings after Certain Syntactic Transformations Apply," in E. Keenan, ed., Formal Semantics, Cambridge University Press, Cambridge.

Emonds, J. E. (1976) A Transformational Approach to English Syntax: Root, Structure Preserving, and Local Transformations, Academic Press, New York.

Emonds, J. E. (1979) "Appositive Relatives Have No Properties," Linguistic Inquiry 10, 2.

Emonds, J. E. (1978) "The Verbal Complex V′-V in French," Linguistic Inquiry 9, 151-175.

Emonds, J. E. (1980) "Inversion Généralisée NP-α: Marque Distinctive de l'Anglais," Langages, Décembre 1980, 60.

Erreich, A., V. Valian and J. Mayer (1978) "Aspects of a Theory of Language Acquisition," paper presented at Language Acquisition: State of the Art, conference held at the University of Pennsylvania, May 19-22, 1978.

Fay, D. (1978) "Transformations as Mental Operations: A Reply to S. Kuczaj," Journal of Child Language 5, 143-149.

Fiengo, R. (1974) Semantic Conditions on Surface Structure, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Fiengo, R. (1976) "On Trace Theory," Linguistic Inquiry 8, 35-61.

Fillmore, C. (1968) "The Case for Case," in E. Bach and R. T. Harms, eds., Universals in Linguistic Theory, Holt, Rinehart and Winston, New York.

Fischer, S. (1971) The Acquisition of Verb-Particle and Dative Constructions, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Fletcher, P. and M. Garman, eds. (1979) Language Acquisition: Studies in First Language Development, Cambridge University Press, Cambridge.

Fodor, J., T. Bever and M. Garrett (1974) The Psychology of Language, McGraw-Hill, New York.

Fraser, B. (1968) The Verb-Particle Construction of English, MIT Press, Cambridge, Massachusetts.

Friedin, R. and H. Lasnik (1981) "Disjoint Reference and WH-Trace," Linguistic Inquiry 12, 1.

Ferguson, C. A. and D. I. Slobin, eds. (1973) Studies of Child Language Development, Holt, Rinehart and Winston, New York.

Gleitman, L. R., H. Gleitman and E. F. Shipley (1972) "The Emergence of the Child as Grammarian," Cognition 1, 137-164.

Gold, E. M. (1967) "Language Identfication in the Limit," Information and Control 10, 447-474.

Goodluck, H. (1978) Linguistic Principles in Children's Grammar of Complement Interpretation, Doctoral Dissertation, University of Massachusetts, Amherst.

Goodluck, H. and T. Roeper (1978) "The Acquisition of Perception Verb Complements," in G. Goodluck and L. Solan, eds., Papers in the Structure and Function of Child Language, University of Massachusetts Occasional Papers in Linguistics 4.

Goodluck, H. and L. Solan (1979) "A Reformulation of the Basic Operations Hypothesis," Cognition 7, 85-91.

Greenberg, J. H., ed. (1966) Universals of Language, second edition, MIT Press, Cambridge, Massachusetts.

Gruber, J. S. (1969) "Topicalization in Child Language," in D. A. Reibel and S. A. Schane, eds., Modern Studies in English, Prentice-Hall, Englewood Cliffs, New Jersey.

Hamburger, H. (1980) "A Deletion Ahead of Its Time," Cognition 8, 389-416.

Hamburger, H. and S. Crain (1979) "Relative Acquisition," to appear in S. Kuczaj, ed., Language Development: Syntax and Semantics, Lawrence Erlbaum Associates.

Hamburger, H. and K. Wexler (1973) "Identifiability of a Class of Transformational Grammars," in K. J. J. Hintikka, J. M. E. Moravcsik and P. Suppes, eds., Approaches to Natural Language, D. Reidel, Dordrecht, Holland.

Hamburger, H. and K. Wexler (1975) "A Mathematical Theory of Learning Transformational Grammar," Journal of Mathematical Psychology 12, 137-177.

Hendrick, R. (1980) "Reduced Questions and the Theoretical Implications of Their Assymetries," unpublished paper, Department of Linguistics, University of North Carolina, Chapel Hill.

Hornby, P. and W. Hass (1970) "Use of Contrastive Stress by Preschool Children," Journal of Speech and Hearing Research 13, 395-399.

Hornstein, N. (1977) "S and $\bar{X}$ Syntax," Linguistic Analysis 3, 2.

Hurford, J. (1975) "A Child and the English Question Formation Rule," Journal of Child Language 2, 299-301.

Jackendoff, R. (1972) Semantic Interpretation in Generative Grammar, MIT Press, Cambridge, Massachusetts.

Jackendoff, R. (1977) $\bar{X}$ Syntax: A Study of Phrase Structure. MIT Press, Cambridge, Massachusetts.

Kaplan, E. and G. Kaplan (1971) "The Prelinguistic Child," in J. Eliot, ed., Human Development and Cognitive Processes, Holt, Rinehart and Winston, New York.

Kayne, R. S. (1975) French Syntax: The Transformational Cycle. MIT Press, Cambridge, Massachusetts.

Klein, S. (1977) "A Base Analysis of the Floating Quantifier in French," in J. A. Kegl, D. Nash and Z. Zaenen, eds., NELS VII, Proceedings.

Klein, S. (1977) Syntactic Theory and a Residual Problem in the Acquisition of Grammar, MS, University of California, Los Angeles.

Klein, S. (1978) "Predicting the Shape of Child Grammars," in Malsch, Hoard and Sloat, eds., Proceedings of the Eighth Annual Meeting of the Western Conference on Linguistics.

Klima, E. S. and U. Bellugi (1966) "Syntactic Regularities in the Speech of Children," in J. Lyons and R. J. Wales, eds., Psycholinguistic Papers, Edinburgh University Press, Edinburgh.

Klima, E. S. and U. Bellugi (1973) "Syntactic Regularities in the Speech of Children," in T. A. Ferguson and D. I. Slobin, eds., Studies of Child Language Development, Holt, Rinehart and Winston, New York. (Later version of 1966 paper.)

Koster, J. (1978) "Why Subject Sentences Don't Exist," in S. J. Keyser, ed., Recent Transformational Studies in European Languages, Linguistic Inquiry Monograph 3, MIT Press, Cambridge, Massachusetts.

Kuczaj, S. (1976) "Arguments against Hurford's 'AUX Copying Rule,'" Journal of Child Language 3, 423-427.

Labov, W. and T. Labov (1978) "Learning the Syntax of Questions," paper presented at Language Acquisition: State of the Art, conference held at the University of Pennsylvania, May 19-22, 1978.

Lapointe, S. (1980) "The Nonnecessity and Acquisition of the Category AUX," unpublished paper, University of California, Irvine.

Lees, R. (1960) The Grammar of English Nominalizations, Mouton, The Hague.

Liberman, M. (1974) "On Conditioning the Rule of Subject-Auxiliary Inversion," in E. Kaisse and J. Hankamer, eds., papers from NELS V. Harvard University, Cambridge, Massachusetts.

Lust, B. (1978) "Constraints on Anaphora in Child Language," paper presented at Language Acquisition: State of the Art, conference held at the University of Pennsylvania, May 19-22, 1978.

McNeill, D. (1966) "Developmental Psycholinguistics," in F. Smith and G. Miller, eds., The Genesis of Language, A Psycholinguistic Approach, MIT Press, Cambridge, Massachusetts.

McNeill, D. (1970) The Acquisition of Language: The Study of Developmental Linguistics, Harper and Row, New York.

McNeill, D. (1981) "The Capacity for the Ontogenesis of Language," in D. I. Slobin, ed., The Ontogenesis of Grammar, Academic Press, New York.

Matthei, E. (1978) "Children's Interpretation of Sentences Containing Reciprocals," in H. Goodluck and L. Solan, eds., Papers in the Structure and Development of Child Language, 4, University of Massachusetts.

May, R. (1977) The Grammar of Quantification, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Mayer, J., A. Erreich and V. Valian (1978) "Transformations, Basic Operations and Language Acquisition," Cognition 6, 1-13.

Menyuk, P. (1969) Sentences Children Use, MIT Press, Cambridge, Massachusetts.

Miller, W. (1964) "The Acquisition of Grammatical Rules by Children," paper presented at LSA, reprinted in C. Ferguson and D. Slobin, eds., Studies of Child Language Development, Holt, Rinehart and Winston, New York.

Milner, J.-C (1976) "Quelques Opérations de détermination en Français," unpublished paper, Université de Lille, III, Lille, France.

Newport, E., L. Gleitman and R. Gleitman (1977) "Mother, Please, I'd Rather Do It Myself: Some Effects and Non-effects of Maternal Speech Style," in C. Snow and C. Ferguson, eds., Talking to Children: Language Input and Acquisition, Cambridge University Press, Cambridge.

Otsu, Y. (1980) "Subjacency Condition and Syntactic Development in Children," abstract of paper presented at GLOW, December 1980.

Otsu, Y. (1981) "Opacity Condition and Syntactic Development in Children," paper presented at NELS XI in Burke and Pustejovsky, eds., Proceedings of the XIth NELS Meeting.

Peters, A. (1977) "Language Learning Strategies," Language 53, 3.

Pierce, C. (1966) Selected Writings, ed. P. P. Wiener, Dover Press, New York.

Pinker, Stephen (1979) "Format Models of Language Learning," Cognition 7, 217-283.

Randeri, M. G. (1980) The Use of Clitics in Spanish between the Ages of Five and Nine, Doctoral dissertation, University of California, Los Angeles.

Reinhart, T. (1976) The Syntactic Domain of Anaphora, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Reinhart, T. (1978) "A Second COMP Position," MS, MIT, Cambridge, Massachusetts.

Riemsdijk, H. van (1978) A Case Study in Syntactic Markedness, Foris, Dordrecht.

Rivas, A. (1977) A Theory of Clitics, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Rizzi, L. (1977) "Violation of the WH-Island Condition in Italian and the Subjacency Condition," unpublished paper, Scuola Normale Superiore, Pisa.

Rochemont, M. (1978) A Theory of Stylistic Rules in English, Doctoral Dissertation, University of Massachusetts, Amherst.

Roeper, T. (1972) Approaches to a Theory of Language Acquisition with Examples from German Children, Doctoral Dissertation, Harvard University, Cambridge, Massachusetts.

Roeper, T. (1973) "Theoretical Implications of Word Order, Topicalization and Inflections in German Language," in C. Ferguson and D. Slobin, eds., Studies of Child Language Development, Holt, Rinehart and Winston, New York.

Roeper, T. (1978) "Linguistic Universals and the Acquisition of Gerunds," H. Goodluck and L. Dolan, eds., Papers in the Structure and Development of Child Language, University of Massachusetts Occasional Papers in Linguistics, IV.

Roeper, T., S. Lapointe, J. Bing and S. Tavakolian (1981) "A Lexical Approach to Language Acquisition," in S. Tavakolian, ed., Language Acquisition and Linguistic Theory, MIT Press, Cambridge, Massachusetts.

Roeper, T. and E. Matthei (1976) "On the Acquisition of All and Some," Papers and Reports on Child Language Development, Stanford University Committee on Linguistics.

Rouveret, A. and J.-R. Vergnaud (1980) "Specifying Reference to the Subject," Linguistic Inquiry 11, 1.

Ruwet, N. (1978) "Une construction absolue en Français," MS., Departement de linguistique générale, Université de Paris VII.

Safir, K. (1981) "Inflection, Government and Inversion," unpublished paper, MIT, Cambridge, Massachusetts.

Schwartz, A. (1972) "Constraints on Movement Transformations," Journal of Linguistics 8, 35-85.

Selkirk, E. (1977) "Some Remarks on Noun Phrase Structure," in P. W. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Slobin, D. I. (1973) "Cognitive Prerequisites for the Development of Grammar," in C. Ferguson and D. Slobin, eds., Studies of Child Language Development, Holt, Rinehart and Winston, New York.

Slobin, D. I. (1978) "Universals and Particulars in the Acquisition of Language," paper presented at the workshop of "Language Acquisition: State of the Art," University of Pennsylvania, May 19-22, 1978.

Smith, N. (1978) The Acquisition of Phonology, A Case Study, Cambridge University Press, Cambridge.

Solan, L. (1978) Anaphora in Child Language, Doctoral Dissertation, University of Massachusetts, Amherst.

Solan, L. and T. Roeper (1978) "Children's Use of Syntactic Structure in Interpreting Relative Clauses," in H. Goodluck and L. Solan, eds., Papers in the Structure and Development of Child Language, Occasional Papers in Linguistics 4, University of Massachusetts, Amherst.

Strozer, J. (1976) Clitics in Spanish, Doctoral Dissertation, University of California, Los Angeles.

Tavakolian, S. (1974) "Contrastive Stress Patterns in Four Year Olds," MS., University of Massachusetts, Amherst.

Tavakolian, S. (1977) Structural Principles in the Acquisition of Complex Sentences, Doctoral Dissertation, University of Massachusetts, Amherst.

Thiersch, C. (1978) Topics in German Syntax, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Vygotsky, L. S. (1962) Thought and Language, MIT Press, Cambridge, Massachusetts.

Wasow, T. (1977) "Transformations and the Lexicon," in P. W. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York.

Wexler, K., P. W. Culicover and H. Hamburger (1975) "Learning Theoretic Foundations of Linguistic Universals," Theoretical Linguistics 2, 213-253.

Wexler, K., and P. W. Culicover (1980) Formal Principles of Language Acquisition, MIT Press, Cambridge, Massachusetts.

White, L. (1977) "Some Proposals for a 'Narrow' Theory of Language Acquisition and Its Interaction with a Theory of Grammar," Montreal Working Papers in Linguistics, No. 9.

Wilkins, N. (1980) "Constraints on Rule Form—Implications for Degree-2 Learnability," School of Social Science Reports, R67, University of California, Irvine.

Williams, E. (1974) Rule Ordering in Syntax, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Williams, E. (1980) "A Readjustment in the Learnability Assumptions," paper presented at workshop on the Logical Problem of Language Acquisition, University of Texas. To appear in C. L. Baker and J. McCarthy, eds., The Logical Problem of Language Acquisition, in press, MIT Press, Cambridge, Massachusetts.